How to Build a $1000 RAC

For Oracle9*i* Real Application Clustering

V.P. of Research and Development Open Systems, Tom Callaghan outlines his customized RAC hardware configuration demonstrated at OracleWorld '02 in San Francisco.

Introduction

Many companies are interested in the benefits of Oracle-9*i* Real Application Clustering. Some of the main benefits include fault tolerance, load balancing, and the ease of upgrading capacity. The hardware required for a production RAC can be very expensive. In fact, the smallest hardware configurations quoted from first tier hardware vendors can run well over $15K to $20K. The heart of a production RAC environment is usually a Storage Area Network. These devices generally start at $15K.

When RAC was introduced, we at Bradmark knew that we needed to provide monitoring and management tools for this technology. However, we did not need to configure a full-scale production RAC with a production RAC cost. We needed the minimum configuration we could find for development and testing.
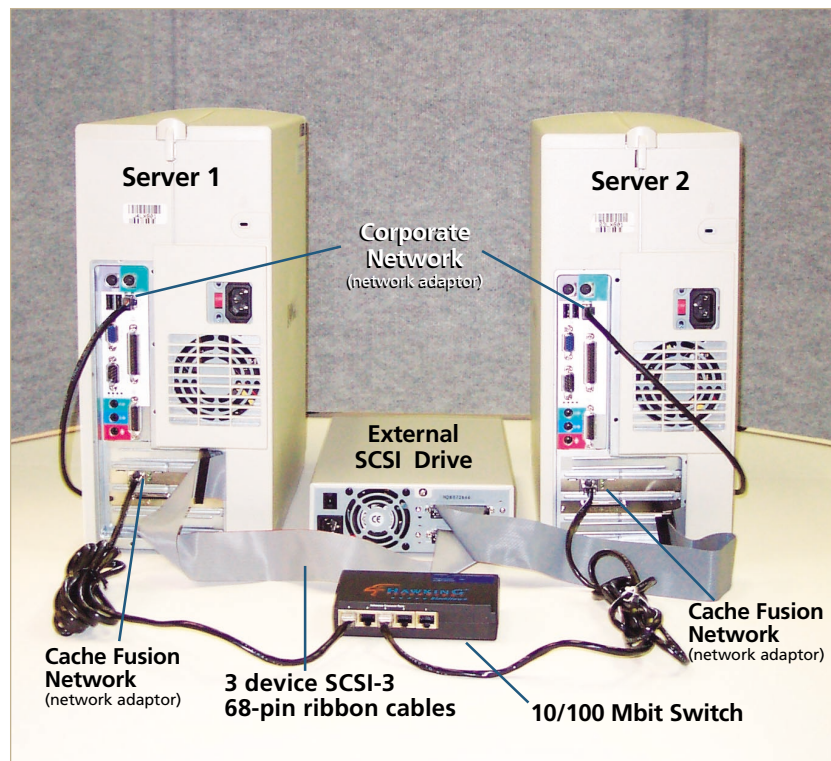
At OracleWorld 2002 we found that many DBAs are in the same boat. They know that RAC is in their future. They want to learn about the technology, hands-on, today, without having to get approval to spend thousands of dollars.

For that purpose we are publishing this paper on the $1000 RAC. Well - its actually the $1098 RAC. I want to note that this cost assumes development licenses for OS and Oracle are already available to the user.

The Concept

I want to emphasize that this configuration **should never run a production application**. Our objective was simply to find the cheapest way we could to configure a RAC. Please listen to and follow the directions of Oracle and your hardware manufacturer when configuring a real RAC for production use.

First, we found two low end desktop computers that we could abscond with from our IT department's spare parts inventory . They both had Celeron 600MHz processors, 256MB RAM and 10GB hard disks.

Server 1
Server 2
Corporate Network (network adapter)
External SCSI Drive
Cache Fusion Network (network adaptor)
Cache Fusion Network (network adaptor)
3 device SCSI-3 68-pin ribbon cables
10/100 Mbit Switch

Custom 9iRAC Cofiguration

bradmark technologies

We decided that a shared SCSI channel with one 18GB disk drive would be adequate for our shared storage.

We decided also to use internal ribbon cables instead of external cables to connect the SCSI bus even though the disk drive was in an external cabinet. The cables and terminators for this configuration were hundreds of dollars cheaper this way.

The computers came with one NIC in each, we had to buy additional NICs and a 5 port 100MB data switch for the cache fusion network. We decided to use the data switch for cache fusion instead of a crossover CAT5 cable because we found that if we took one of the PCs down, the other PC would detect "cable unplugged", and thus the cache fusion network became invalid.

We installed an Oracle 9.2 RAC with the shared disk on the SCSI bus. We used the Oracle File System (OFS) to access the shared disk. We used Oracle Cluster Manager (CM) to create a RAC cluster (Microsoft Clustering is something different and not used here). This is the standard, default configuration for RAC on Windows.

The Parts List

Description	Unit Price	Qty	Extended Price
DELL Dimension 600CX computers with 600MHz Celeron processors, 256MB RAM, 10GB hard disk	Approx Value $300	2	$600
Adaptec ASC-29160LP U160 SCSI PCI LVM controllers	$99	2	$198
Quantum Atlas 10k3_18_WLS 18GB SCSI disk drive	$100	1	$100
Generic external SCSI disk drive cabinet with SCSI-3 68 pin connectors	$74	1	$74
Cables Unlimited CUS-F-005 34" 3 device SCSI-3 68 pin ribbon cables	$20	2	$40
Cables Unlimited SCS-5480 SCSI-3 terminators HDB68F *(internal, active, plugs into M connector)*	$13	2	$26
Hawkings PN105ES 5 port 10/100MB data switch with one 100MB NIC included	$20	2	$40
3' CAT5 cables	5	4	$20
		Total:	**$1,098**

The Assembly

Prior to installing any of the new hardware, we first performed clean installs of MS Windows 2000 Advanced Server on our two PCs. After a clean install, we shut the computers down for hardware installation. Our plan is to install one of these for LINUX. At that time, we'll update the paper. However, we wanted to get this paper out to you ASAP.

Physically install the SCSI cards into the two computers. The default SCSI ID for the cards is 7. Later we will reconfigure one to 6.

Install the SCSI disk drive in the external SCSI cabinet. By default, our disk drive was set to SCSI ID 2. We just left it there.

Each of the SCSI cables has 3 connectors. On the end with two of the connectors close together, install one terminator on the end connector, the internal connector of the SCSI card on the middle connector, and the SCSI disk drive cabinet on the far end connector. Do this for both computers. When you are done, the systems should look something like this. (Insert picture)

Install the NICs for the cache fusion network. Connect the cache fusion network with the CAT5 cables and one data switch.

The second data switch is to provide for a standalone environment. I am assuming you already had a NIC in each computer. Connect them together with CAT5 cables and the second data switch. It would make sense to have another client computer on the network as well. Alternatively, you can omit the second data switch and simply connect each computer to your existing intranet.

We are now ready to configure the hardware.

Hardware Configuration

It may be helpful to disable the Dell logo screen, so that you can see computer startup messages.

Go into the SCSI card BIOS Setup on each computer and make the following changes.

✔ **Set the SCSI IDs.** Set one card at SCSI ID 7 and the other card at SCSI ID 6.

✔ **Disable on-card termination.** Set the "Host Adapter SCSI Termination" option to "disabled". If it is set to "automatic" (the default), the on-card terminator may erroneously be activated. We want to use the terminators we have physically placed on the SCSI cables (which are at the ends of the SCSI bus).

✔ **Disable resetting of the SCSI bus / hard drive on reboot.** In the column for SCSI ID 2 (the hard drive) in "SCSI Device Configuration", set "Enable Disconnection" to "no" and "Send Start Unit Command" to "no". Under "Advanced Configuration Options", set "Reset SCSI Bus at IC Initialization" to "disabled". Disabling all of these options ensures the SCSI bus and shared drive are not reset when one PC reboots. **Set miscellaneous options.** Set the "SCSI Parity Checking" option to "disabled". Optionally, disable the "Host Adapter BIOS" under "Advanced Configuration Options" if it is not needed, to ensure no complications.

In Windows, set up the Network Connections. The NICs on the cache fusion (private) network should be given private addresses, such as 192.168.1.1 and 192.168.1.2. The NICs on the stand-alone (public) network, or your intranet, will have IP addresses for another subnet, such as 192.168.2.1 and 192.168.2.2. Enable TCP/IP and "File and Printer Sharing" on all of the Network Connections.

Software Installation

Add the following entries into the \winnt\system32\drivers\etc\hosts file on both PCs:
192.168.1.1   privA privA.domain.com
192.168.1.2   privB privB.domain.com
192.168.2.1   pubA pubA.domain.com
192.168.2.2   pubB pubB.domain.com

where 192.168.1.1, 192.168.1.2 are the private network addresses, 192.168.2.1, 192.168.2.2 are the public network addresses, "domain.com" is your stand-alone or intranet domain name, and "A", "B" are your PC names. These entries are needed to distinguish between the private and public networks.

All of the following assumes you are installing Oracle version 9.2.

Software Installation
Continued

The Oracle 9iR2 RAC installation will, when we are done, install several Windows NT Services. Briefly, they are:

OracleClusterVolumeService – Oracle File System.
This gives us access to the shared disk.

OracleCMService9i – Oracle Cluster Manager. This service, among other things, detects other PCs in the RAC and decides which PC is the boss.

OracleGSDService – This service coordinates tasks among PCs in the cluster.
This service depends on the CM Service.

OracleOraHome92RACTNSListenerlistener_publicA, _publicB – Standard TNS listener service for computer "A" or "B", respectively (with Oracle Home""OraHome92RAC"). This service depends on the GSD Service.

OracleServiceA, OracleServiceB"– Standard instance service for computer "A" and""B", respectively. This service depends on the GSD and CM Services.

Now let's begin. We assume that we will conduct the installation on machine A. You could choose B instead.

Log into machine A with a domain administrator account. This account must be able to access the hidden "C$" share on both machines. Establish this access by opening a command prompt and running:

NET USE \\A\C$ password /USER:mydomain\administrator
NET USE \\B\C$ password /USER:mydomain\administrator

where
"**mydomain**" is the Windows domain and "administrator", "**password**" are appropriate.

On the Oracle 9iR2 Cluster File System (CFS) CD, follow the instructions in \doc\html\ocfs_relnotes.htm to install the Cluster File System, Cluster Components, and Oracle 9iR2 Enterprise Edition. This involves:

- Running the Oracle Cluster Setup Wizard (clustersetup.exe). Choose private interconnect, listing the machine names privA and privB as specified in the hosts file. Put both the Oracle Home and the data files under CFS control.

- Running \install\win32\setup.exe on the CFS CD to install Cluster Components using the Universal Installer.

- Running Universal Installer for Oracle 9.2 Enterprise Edition. Note that you will **not** be creating a database at this time.

- Patching the Oracle 9.2 installation with ORACLE_HOME\cfspatch. There are instructions in several text files in that directory. Note that the instructions have several obvious typos with regard to filenames, so be sure to use every file in the cfspatch directory. At the end of the instructions, you will be starting the GSD service from a command prompt (after installation the GSD service will always be started automatically on boot).

- Configuring the listener. Note that the instructions will tell you to disable the default listener and a new one will be created implicitly.

- Creating the database. Be sure to run dbca from a command prompt.

Software Installation
Continued

By default the Oracle installation added dependencies between several of the NT services it installed. However, for robustness, you should add a dependency between the Listener service and the GSD service. You will need to add the following registry entry to each machine using REGEDT32 and reboot both machines:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\OraHome92RACListenerlistener_publicA (or __publicB)\DependOnService, REG_MULTI_SZ, OracleGSDService

Verify that the instance service (OracleServiceA, for example) depends on the GSD and CM services.

Reboot the machines to test whether the NT services work properly. First the Object and ClusterVolume services start automatically, then CM, then GSD, then the listener and the instance.

After rebooting check \WINNT\SYSTEM32\osd9i\cmsrvr_publicA.log (or appropriate) to make sure both nodes of the cluster are recognized. It should say "2 active node(s)".

The Oracle 9.2 RAC is now ready for use.

Troubleshooting

If one PC reboots, the other one freezes or gets a BSOD. This indicates that the one or both cards are resetting the SCSI bus / hard drive on reboot. Be sure to set "Enable Disconnection" and "Send Start Unit Command" to "no" for SCSI ID 2 (the hard drive) on both cards.

You encounter long timeouts when browsing the shared drive. This indicates that the termination on the SCSI bus is incorrect. Make sure the on-card termination on both SCSI cards is disabled, and ensure the terminators are placed on the far ends of the SCSI bus (the ends of the cables closest to the cards). The terminators may be malfunctioning or have the wrong resistance for your needs.

Installation of Oracle 9iR2 Enterprise Edition fails. This can be a result of incorrect termination. See "You encounter long timeouts when browsing the shared drive".

About Tom Callaghan
Vice President Research and Development Open Systems, Mr. Callaghan joined Bradmark through our acquisition of Pace Systems in June 2000. His work in Information Technology began in 1982, and since 1991 his focus has been on the database solutions market. Mr. Callaghan received his Bachelor of Arts in Computer Information Systems from the University of Houston Clear Lake in 1982 and soon after joined Compaq Computers. While at Compaq, he managed Technical Support for the corporate datacenters worldwide. In this capacity he ran many successful multi-million dollar projects.

About Bradmark
For 20 years, Bradmark has provided a business advantage by ensuring that the databases, which are vital to the organization's competitive advantage, meet or exceed health and availability requirements.

Bradmark is committed to providing data management solutions in support of our customers' business challenges.

To order or for more information on other Bradmark products:
Phone: **(800) 621-2808** or
Outside the U.S.: (713) 621-2808
Fax: (713) 621-1639
Or, visit our Web site: **www.bradmark.com**

In the UK:
Bradmark Technologies UK Ltd.
Tel: +44 (0)870 240 6285
Fax: +44 (0)121 321 3555

In Germany:
Bradmark Deutschland GmbH
Tel: +49 (0)89 962 09012
Fax: +49 (0)89 962 80860