



**Technical report:**  
**Oracle10g RAC with AIX and IBM System  
Storage N series**

*Best-Practices Installation*

• • • • • • • • •

*Document NS3479-0*

August 22, 2007



## Table of contents

<b>Abstract</b> .....	<b>1</b>
<b>Introduction</b> .....	<b>1</b>
<b>Assumptions</b> .....	<b>1</b>
<b>The Server/System Environment</b> .....	<b>2</b>
<b>Requirements</b> .....	<b>3</b>
Hardware Used .....	3
Cluster Nodes.....	3
Storage Infrastructure.....	3
Software Used .....	3
<b>Setup for IBM System Storage N series</b> .....	<b>3</b>
<b>Patches, Environment and Operating System Settings</b> .....	<b>4</b>
Patches and Filesets.....	4
Operating System Settings .....	5
Update UDP and TCP packet settings .....	6
<b>Preinstall Setup Tasks (Cluster Nodes)</b> .....	<b>7</b>
<b>Installation Procedure</b> .....	<b>10</b>
<b>Appendix</b> .....	<b>10</b>
<b>Trademarks and Special Notices</b> .....	<b>11</b>



## Abstract

---

*This technical report covers the best-practice, certified-configuration installation of IBM System Storage N series in an NFS environment for IBM AIX 5L on the IBM System p platform. The System p models are running the AIX 5L Version 5.3 Technology Level-4 (AIX 5.3TL-04) operating system. The database is Oracle10g Real Application Clusters Release 2 with Oracle Cluster Ready Services.*

## Introduction

---

This technical report covers the installation of IBM<sup>®</sup> System Storage<sup>™</sup> N series in a network file system (NFS) environment for IBM AIX on the IBM System p<sup>™</sup> platform. The System p models are running the AIX 5L Version 5.3 Technology Level-4 (AIX 5.3TL-04) operating system. The database is Oracle10g<sup>™</sup> Real Application Clusters (RAC) Release 2 (hereafter referred to as Oracle10g R2 RAC) with Oracle Cluster Ready Services (hereafter referred to as Oracle CRS). This is now a certified configuration and, as such, the components presented in this paper have to be used in the same combination to gain support from all parties involved. The only exception to this is the application of certain patches (as defined and required by all the vendors in this configuration). This document will also cover the patches and recommendations for running Oracle10g R2 RAC on IBM N series storage in NFS environments.

## Assumptions

---

We assume that readers are familiar with Oracle10g R2 RAC and the operation of IBM System Storage N series. We also assume that readers are familiar with the operation of the IBM AIX<sup>®</sup> environment and installation of Oracle<sup>®</sup> patches and any relevant AIX patches. It is also important to be familiar with all networking terminology and implementations.

## The Server/System Environment

The configuration presented in this document is based on the Oracle10g R2 RAC certification environment specified by Oracle and IBM N series.

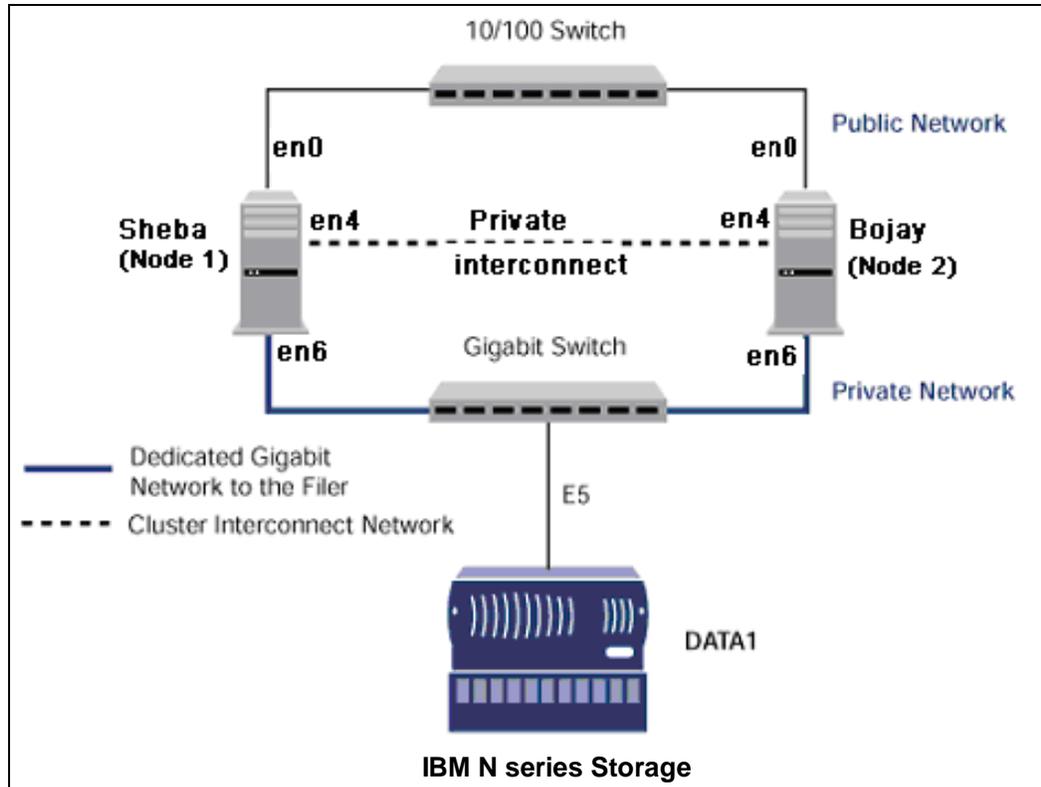


Figure 1. Oracle10g R2 RAC on IBM pSeries PowerPC servers with IBM N series storage.

Figure 1 illustrates a typical configuration of Oracle10g R2 RAC with IBM N series storage and IBM pSeries systems with the IBM PowerPC™ processor, running AIX 5L Version 5.3. This is a scalable configuration and allows users to scale horizontally and internally in terms of processor, memory, and storage.

As shown in the network diagram, we recommend that you dedicate a private network connection between the Oracle10g R2 RAC servers and the IBM N series storage. This is accomplished using a dedicated gigabit network (with a gigabit switch) to the IBM N series storage. A dedicated network connection is beneficial for the following reasons:

- In an Oracle10g R2 RAC environment, it is important to eliminate contentions and latencies.
- Providing a separate network ensures security.

The cluster interconnect is an essential part for 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 Used

#### Cluster Nodes

- Two IBM pSeries 610 models—Model 6E1, 64-bit, PowerPC
- 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 System Storage N series).

#### Storage Infrastructure

- One IBM N series storage system with Data ONTAP® 7.1 or later
- One gigabit switch with at least four ports
- One gigabit network interface card (NIC) in the system
- One or more disk shelves, based on the disk space requirements.

### Software Used

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

- AIX 5.3 TL04 (oslevel -s gives 5300-04-1)
- Oracle10g R2 (10.2.0.1) RAC license.

## Setup for IBM System Storage N series

---

1. Please configure an IBM N series storage system running Data ONTAP 7.1 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 three volumes on the storage (Data1) as listed below.

volhome	Shared Oracle and CRS Home(Binaries)
voldata	Oracle datafiles, control files and CRS files
vollog	database logs, a copy of control file and archive logs.

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

```
Data1> vol create voldata 14
```

**Note:** We created volume voldata with 14 disks and volumes vollog and volhome with 8 disks each. You can create your volumes based on your workload needs.

Edit the /etc/exports file on IBM N series storage (Data1) and add the following entries to that file:

```
/vol/volhome -anon=0  
/vol/voldata -anon=0  
/vol/vollog -anon=0
```



- b. Execute the following command at the storage system console:

```
Data1> 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 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 and your setup may vary depending upon requirements.

## Patches, Environment and Operating System Settings

### Patches and Filesets

Before your Oracle10g R2 RAC install, the following patches/filesets need to be applied on IBM systems. Some of these patches/filesets may already be applied to your system. Please verify if they already exist before applying them.

To determine whether the required filesets are already installed and committed, enter a command similar to the following:

```
# lslpp -l bos.adt.base
```

If a fileset is not installed and committed, then install it. Here is a list of required filesets:

- bos.adt.base
- bos.adt.lib
- bos.adt.libm
- bos.cifs\_fs.rte (if not installed, OUI will issue a warning during the pre-install check)
- bos.perf.libperfstat
- bos.perf.perfstat
- bos.perf.proctools
- rsct.basic.rte (for RAC configurations only)
- rsct.compat.clients.rte (for RAC configurations only)
- xIC.aix50.rte:7.0.0.4 or higher (see the following note)
- xIC.rte:7.0.0.1 or higher (see the following note).

The xIC filesets are provided by the installation of the “IBM XL C/C++ Enterprise Edition V7.0” compiler or the “IBM XL C/C++ Enterprise Edition V7.0 for AIX Runtime Environment Component.” Note that the “IBM VisualAge® C++ V6.0 Runtime Environment for AIX” is installed by default since it is packaged with the current distributions of AIX 5L V5.3 and V5.2. Therefore, Oracle10g R2 users who do not install the “IBM XL C/C++ Enterprise Edition V7.0” compiler must install the “IBM XL C/C++ Enterprise Edition V7.0 for AIX Runtime Environment Component.” This can be easily downloaded with no license requirement from: [ibm.com/support/docview.wss?uid=swg24009788](http://ibm.com/support/docview.wss?uid=swg24009788).



Next, verify and install PTF (program temporary fix) for APAR IY76140. To determine whether an APAR is installed, enter a command similar to the following:

```
# /usr/sbin/instfix -i -k "IY76140"
```

If an APAR is not installed, download it from the following Web site and install it:

<https://techsupport.services.ibm.com/server/aix.fdc>

## Operating System Settings

On AIX systems, the default ulimits for individual users are set in /etc/security/limits. As a root user, add the following entries for the root and Oracle users:

root:

```
fsize = -1
core = 2097151
cpu = -1
data = -1
rss = -1
stack = -1
nofiles = -1
```

Oracle:

```
fsize = -1
core = 2097151
cpu = -1
data = -1
rss = -1
stack = -1
nofiles = -1
```

This needs to be done on all nodes of the cluster. A server reboot might be necessary to activate updated limits. After you modify the settings, “ulimit -a” command should display the following:

```
# ulimit -a

time(seconds) unlimited
file(blocks) unlimited
data(kbytes) unlimited
stack(kbytes) 2097152
memory(kbytes) unlimited
coredump(blocks) 2097151
nofiles(descriptors) unlimited
```

Verify the same for the Oracle user. Next, you need to increase Oracle user processes to 8192. You can increase the processes for the Oracle user via the “smit” utility on AIX based systems.



## Update UDP and TCP packet settings

Oracle RAC uses User Datagram Protocol (UDP) for interprocess communications. You may need to tune UDP and transmission control protocol (TCP) kernel settings to improve Oracle performance. On AIX systems, the TCP and UDP specific settings can be verified by the command 'no -a'. Please change/verify the values of the TCP/UDP parameters as indicated below.

```
tcp_sendspace=262144
tcp_recvspace=655360
udp_sendspace=65536
udp_recvspace=262144
rfc1323=1 (For any TCP/ UDP parameters settings of over 64K.)
```

These parameters can be set with the “smit” utility on AIX-based servers. Make sure that these parameter values are retained during subsequent reboots. These parameters may also be set via the command line as a root user. However, they will need to be set again after every reboot or can be placed in the file /etc/rc.net. For example,

```
# no -o rfc1323=1
```

**Note:** The value of the udp\_recvspace parameter should be at least 10 times the value of the udp\_sendspace parameter because UDP might not be able to send a packet to an application before another packet arrives.

These settings are generally recommended for Oracle10g R2 RAC functionality.



## Preinstall Setup Tasks (Cluster Nodes)

---

This information is intended for both nodes in the participating cluster unless specified otherwise.

1. Have two AIX 5L System p servers ready with the latest recommended patches/Filesets and operating system settings, as discussed in Section 6.
2. Install/configure NICs in the cluster nodes (three per node).
  - a. Public IP: As indicated by name.
  - b. Private interconnect: Connect one gigabit NIC back to back to the other node for cluster interconnects.
  - c. Server connection to IBM N series storage: Connect one gigabit NIC to the gigabit switch, which will connect to the gigabit NIC on the IBM N series storage.
3. Configure the network interfaces on each node.
  - a. Configure the three network interfaces as indicated below.

```
# sheba (Host 1)
```

```
en0 - ip: 10.61.161.57, netmask 255.255.255.0  
en6 - ip: 10.61.169.57, netmask: 255.255.255.0  
en4 - ip: 10.0.0.1, netmask: 255.255.255.0
```

```
# bojay (Host 2)
```

```
en0 - ip: 10.61.161.56, netmask: 255.255.255.0  
en6 - ip: 10.61.169.56, netmask: 255.255.255.0  
en4 - ip: 10.0.0.2, netmask: 255.255.255.0
```

Where:

Interface en0 is the public ip for each node.

Interface en6 on both cluster nodes is connected to the gigabit switch for storage I/O.

Interface en4 on both cluster nodes is connected back-to-back for cluster private interconnects.



- b. Update the /etc/hosts file on the cluster nodes and add entries for public, private, and virtual IP (VIP) addresses. Please note that in addition to the preconfigured public and private network, Oracle10g requires additional IP addresses that will be mapped to the public address as VIPs. If a node fails when an application or user makes a connection using a VIP, the Oracle clusterware will transfer the VIP address to another surviving instance. You should add the VIP to the /etc/hosts file on all nodes in the cluster as well as all nodes accessing the database.

A sample for /etc/hosts entries

```
# Internet Address Hostname      # Comments
10.61.161.57 sheba        sheba.rtp.ibm.com
10.61.169.57 sheba-en6
10.0.0.1    sheba-priv  sheba-priv.rtp.ibm.com
10.61.161.71 sheba-vip   sheba-vip.rtp.ibm.com

10.61.161.56 bojay        bojay.rtp.ibm.com
10.61.169.56 bojay-en6
10.0.0.2    bojay-priv  bojay-priv.rtp.ibm.com
10.61.161.72 bojay-vip   bojay-vip.rtp.ibm.com
```

- 4. Ensure the connectivity of each interface via the ping command. (interconnects, public IPs, and storage)

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/filesystems` file on all server nodes and add the following entries:

```

/voldata:
    dev          = "/vol/Voldata"
    vfs          = nfs
    nodename     = Data1
    mount        = true
    options      = cio,rw,bg,hard,nointr,proto=tcp,vers=3,
noac,rsize=32768,wsiz=32768,timeo=600
    account      = false
/volhome:
    dev          = "/vol/Volhome"
    vfs          = nfs
    nodename     = Data1
    mount        = true
    options      = rw,bg,hard,nointr,proto=tcp,vers=3,
rsiz=32768,wsiz=32768,timeo=600
    account      = false
/vollog:
    dev          = "/vol/Vollog"
    vfs          = nfs
    nodename     = Data1
    mount        = true
    options      = cio,rw,bg,hard,nointr,proto=tcp,vers=3,
noac,rsiz=32768,wsiz=32768,timeo=600
    account      = false
/quorum:
    dev          = "/vol/Voldata/quorum"
    vfs          = nfs
    nodename     = Data1
    mount        = true
    options      = cio,rw,bg,hard,intr,proto=tcp,vers=3,
noac,rsiz=32768,wsiz=32768,timeo=600
    account      = false

```

Where:

- `Data1` is the name of the IBM N series storage system.
- `voldb`, `vollog`, `volhome`, and `quorum` are the mount points on the cluster nodes.
- `Quorum` is a separate mount point for CRS files. CRS files (cluster registry file and voting disk file) can reside in the same `/voldata` volume, but must be mounted with the “intr” mount option; this option is required in failure scenarios in which CRS has to evict a node.
- During the Oracle CRS install, please make sure to indicate a path starting with `/quorum` directory when prompted for the Cluster registry file (`ocr`) and voting disk (`css`) file location.
- **Note:** “cio” mount option should not be used for `ORACLE_HOME` and `CRS_HOME` (the volume where CRS and Oracle binaries are placed).
- Oracle install will fail if you dynamically mount the NFS volumes without adding entries in `/etc/filesystems` file.



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

```
#mkdir /voldb
#mkdir /vollog
#mkdir /volhome
#mkdir /quorum
```

Mount the exported volumes on the mount points created (above) on all cluster nodes. It is always a good idea to verify the mount options by using the mount command on each node. After mounting the NFS volumes, change the ownership of these mounted volumes to Oracle user.

## Installation Procedure

---

For an Oracle10g R2 RAC install on AIX, please refer to Oracle install documentation (part #B14201-03). The following link provides instructions for Oracle10g R2 RAC install on AIX:

[http://download-west.oracle.com/docs/cd/B19306\\_01/install.102/b14201/toc.htm](http://download-west.oracle.com/docs/cd/B19306_01/install.102/b14201/toc.htm).

## Appendix

---

### Sample .profile file for the Oracle user:

```
export ORACLE_BASE=/volhome/ora10g;
export ORACLE_PRODUCT=$ORACLE_BASE/product;
export ORACLE_HOME=$ORACLE_PRODUCT/10.2.0/db_1;
export ORACLE_CRS=$ORACLE_PRODUCT/10.2.0/crs_1;
export ORACLE_SID=oce;
export
LD_LIBRARY_PATH=$ORACLE_HOME/lib:$ORACLE_CRS/lib:$ORACLE_HOME/lib32:$LD_LIBRAR
Y_PATH;
export LIBPATH=$ORACLE_HOME/lib:$ORACLE_CRS/lib:$ORACLE_HOME/lib32:$LIBPATH
export PATH=$PATH:$ORACLE_HOME/bin:$ORACLE_HOME:$ORACLE_CRS/bin:/usr/java14/bin;
export AIXTHREAD_SCOPE=S;
```



## Trademarks and Special Notices

---

© International Business Machines 1994-2007. IBM, the IBM logo, System Storage, 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.

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

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

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

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

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.