



NetApp®
Go further, faster

Technical Report

Oracle Real Application Clusters 11g Release 1 with AIX 6L Version 6.1 and NetApp Storage

Oracle Alliance Engineering Team, NetApp
June 2010 | TR-3542

ARCHIVAL COPY
Contents may be out-of-date

ABSTRACT

This technical report covers the installation of NetApp® storage in an NFS environment for IBM® AIX® 6L on IBM System p5 520Q, model 9111-520 running AIX 6L Version 6.1 Technology Level 4 (6100-04). The database used is Oracle® 11g Real Application Clusters release 1. The focus of this paper is technical, and the reader should be comfortable with AIX system administration, Oracle Database 11g administration, network connectivity, and NetApp system administration.

TABLE OF CONTENTS

1	INTRODUCTION	3
2	ASSUMPTIONS	3
3	SERVER/SYSTEM ENVIRONMENT	3
4	REQUIREMENTS	4
4.1	HARDWARE REQUIREMENTS	4
4.2	SOFTWARE REQUIREMENTS	5
5	NETAPP STORAGE SETUP	5
6	PATCHES, ENVIRONMENT, AND OS SETTINGS	6
6.1	PATCHES AND FILE SETS	6
6.2	OPERATING SYSTEM SETTINGS	7
7	PREINSTALLATION SETUP (CLUSTER NODES)	9
8	INSTALLATION PROCEDURE	11
9	APPENDIX: SAMPLE PROFILE FILE	11
	REVISION HISTORY	11

ARCHIVAL COPY
Contents may be out-of-date

1 INTRODUCTION

This technical report covers the installation of NetApp storage in the NFS environment for IBM AIX 6L on IBM System p5 520Q, model 9111-520. The IBM System p5 520Q Express platform is a new breed of 4-core powered by IBM's latest Quad-Core Module (QCM) 1.65 GHz POWER5+ processor cores, running AIX 6L version 6.1 technology level 4 (6100-04) operating system. The database is Oracle11g Real Application Clusters release 1 (hereafter referred to as Oracle Database 11g R1 RAC) with Oracle Cluster Ready Services (hereafter referred to as Oracle CRS). This is 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 also covers the patches and recommendations for running Oracle Database 11g R1 RAC on NetApp storage in an NFS environment.

2 ASSUMPTIONS

This report assumes familiarity with:

- Oracle Real Application Clusters 11g release 1 and the operation of NetApp storage systems
- Operation of the IBM AIX 6L environment and installation of Oracle patches and any relevant AIX patches
- Networking terminology and implementations

3 SERVER/SYSTEM ENVIRONMENT

The configuration presented in this document is based on the Oracle Database 11g R1 RAC certification environment specified by Oracle and NetApp.

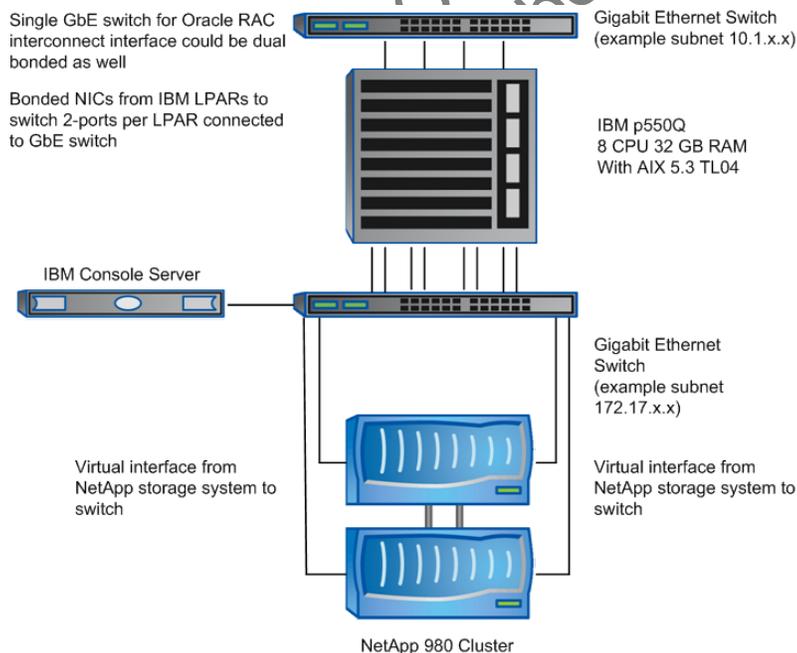


Figure 1) Oracle Database 11g R1 RAC on IBM pSeries POWER5 server with NetApp storage.

Figure 1 illustrates a typical configuration of Oracle Database 11g R1 RAC with NetApp storage and IBM pSeries POWER5 server running AIX 6L Version 6.1. This configuration allows you to scale vertically or horizontally in terms of processor, memory, and storage.

NetApp recommends creating a dedicated private network connection between the Oracle Database 11g R1 RAC servers and the NetApp storage. This is accomplished using a dedicated gigabit network (with a gigabit switch) to the NetApp storage. A dedicated network connection is beneficial for the following reasons:

- In an Oracle Database 11g R1 RAC environment, it is important to eliminate any contentions and latencies.
- A separate network provides security.

The cluster interconnect is an essential part for Oracle Database clusters. Together with cache fusion, the cluster interconnect monitors the heartbeat of the servers in the cluster group.

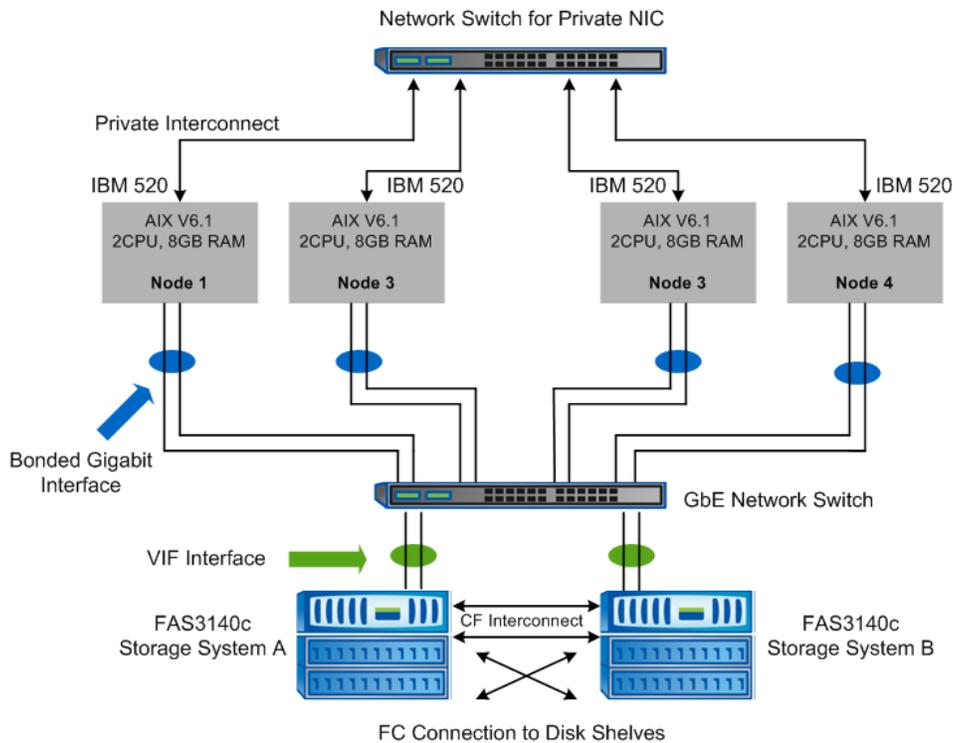


Figure 2) Oracle Database 11g R1 RAC configuration.

4 REQUIREMENTS

4.1 HARDWARE REQUIREMENTS

CLUSTER NODES

- One IBM pSeries 520Q, Model 9111-520, IBM POWER5 Server
- 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 NetApp storage)

STORAGE INFRASTRUCTURE

- One NetApp FAS2xx/F7xx/F8xx/FASF9x/FAS30xx system with Data ONTAP® 7.0 or later
- One gigabit switch with at least four ports
- One gigabit NIC in the system
- One or more disk shelves, based on the disk space requirements

4.2 SOFTWARE REQUIREMENTS

In all the nodes in the participating cluster unless specified otherwise:

- AIX 6.1 TL02 SP1 (6100-02-01)
- Install the latest fixes from <http://www-912.ibm.com/eserver/support/fixes/>.
- Oracle Database 11g R1 (11.1.0.1), with Real Application Clusters license
- Minimum software versions and patches required to support Oracle products on IBM pSeries
<https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&doctype=BULLETIN&id=282036.1>

5 NETAPP STORAGE SETUP

1. Configure a NetApp storage system running Data ONTAP 8.0 with NFS and SnapRestore[®] license keys.
2. Create and export volumes for storing Oracle Database and other necessary files on the storage:
 - i. Create volumes on the storage (Data1):

```
/u01/app/oracle      <Shared home for clusterware & database binaries>
/oradata             <Shared location for DataFiles>
/oractl              <Shared location for ControlFiles>
/oralog              <Shared location for Online Redo Logs & Archived
                    Logs>
/storage             <Shared location for Oracle Cluster registry &
                    voting disks>
```

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

```
Data1> vol create vol1 aggregate_name size
```

Note: We created volume `vol1` under the aggregate `aggregate_name` with size 'size'. You can create your volumes based on your workload needs.

Edit the `/etc/exports` file on NetApp storage (Data1) and add the following entries:

```
/vol/vol1 -anon=0
/vol/vol2 -anon=0
...
```

- ii. Execute the following command on the storage system console:

```
Data1> exports -a
```

Note:

NetApp recommends using flexible volumes in your database environment. NetApp 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. For database layout best practices on NetApp storage, see NetApp [TR 3411](#).

6 PATCHES, ENVIRONMENT, AND OS SETTINGS

6.1 PATCHES AND FILE SETS

Before installing Oracle Database 11g R1 RAC, apply the following patches/file sets on each IBM server. Verify that these patches exist in your system before applying them.

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

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

If a file set is not installed and committed, then install it. Here is a list of required file sets:

- bos.adt.base
- bos.adt.lib
- bos.adt.libm
- bos.cifs_fs.rte (if not installed, OUI will issue a warning during the preinstall 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)
- xlc.aix61.rte:9.0.0.1 or higher
- xlc.rte:9.0.0.1 or higher

Note:

The xlc file sets are provided by the installation of the IBM XL C/C++ Enterprise Edition V9.0 compiler or the IBM XL C/C++ Enterprise Edition V10.1.0 for AIX Runtime Environment Component. Note that the VisualAge C++ V6.0 Runtime Environment for AIX is installed by default since it is packaged with the current distributions of AIX 6L V6.1. Therefore, Oracle Database 11g Release 1 users who do not install the IBM XL C/C++ Enterprise Edition V9.0 compiler must install the IBM XL C/C++ Enterprise Edition V10.1.0 for AIX Runtime Environment Component. This can be easily downloaded with no license requirement from:

<http://www-01.ibm.com/support/docview.wss?uid=swg24019829>

<http://www-01.ibm.com/software/awdtools/xlcpp/aix>

Next, verify and install PTF (program temporary fix) for APAR IZ41855, IZ51456 and IZ52319. To determine whether an APAR is installed, enter a command similar to the following:

```
# /usr/sbin/instfix -i -k "IZ41855"
```

If an APAR is not installed, download and install it from <https://techsupport.services.ibm.com/server/aix.fdc>

6.2 OPERATING SYSTEM SETTINGS

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

```
root:
  fsize = -1
  core = -1
  cpu = -1
  data = 512000
  rss = 512000
  stack = 512000
  nofiles = 2000
```

```
oracle:
  fsize = -1
  core = -1
  cpu = -1
  data = 512000
  rss = 512000
  stack = 512000
  nofiles = 2000
```

2. Perform these steps on all nodes of the cluster. A server reboot is necessary to activate updated limits if changes are made to the file manually. After all the settings are activated running `ulimit -a` command displays the following:

```
# ulimit -a
time(seconds) unlimited
file(blocks) unlimited
data(kbytes) 256000
stack(kbytes) 256000
memory(kbytes) 256000
coredump(blocks) 2097151
nofiles(descriptors) 2000
```

3. Verify the same for the oracle user. Next, increase oracle user processes to 8192 by using the `smit` utility on AIX-based servers. For more information on the `smit` tool, see the Operating System Administration Guide or contact your system administrator.

UPDATE UDP AND TCP PACKET SETTINGS

Oracle Real Application Clusters uses user datagram protocol (UDP) for inter-process communications. Tune the UDP and TCP kernel settings for optimal Oracle performance. On AIX systems, the TCP and UDP specific settings can be verified using the `no -a` command. Change/verify the values of the TCP/UDP parameters as indicated below:

```
tcp_sendspace=262144
```

```
tcp_recvspace=65536
```

```
udp_sendspace=65536
```

```
udp_recvspace=655360
```

```
rfc1323=1 (for any TCP/UDP parameter settings over 64K)
```

These parameters can be set using 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 execute the following command to set the value immediately,

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

Add entries similar to the following to the `/etc/rc.net` file for each parameter that you changed in the previous step:

```
if [ -f /usr/sbin/no ] ; then
    /usr/sbin/no -o udp_sendspace=65536
    /usr/sbin/no -o udp_recvspace=655360
    /usr/sbin/no -o tcp_sendspace=262144
    /usr/sbin/no -o tcp_recvspace=65536
    /usr/sbin/no -o rfc1323=1
    /usr/sbin/no -o sb_max=2*655360
    /usr/sbin/no -o ipqmaxlen=512
fi
```

By adding these lines to the `/etc/rc.net` file, the values persist when the system restarts.

Note:

The value of `udp_recvspace` parameter should be at least 10 times the value of the `udp_sendspace` parameter as UDP may not be able to send a packet to an application before another packet arrives.

These settings are generally recommended for Oracle11g R1 RAC functionality. For additional performance-related guidelines, see [TR 3871: AIX Protocol Performance Comparison with Oracle Database 11g Release 2](#).

In the virtual memory setting, `minperm` was changed to 5%, instead of 20% as in default, to cap non-computational pages.

Note:

Tuning I/O and memory with the `vm tune` command is now replaced by the `vmo` and `ioo` commands, where `ioo` handles all the I/O-related parameters while `vmo` handles the virtual memory management ones.

```
vmo -L
```

7 PREINSTALLATION SETUP (CLUSTER NODES)

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

1. Check that all four LPARs of the IBM System p5 520Q, model 9111-520 server have the latest recommended patches/file sets and OS settings as discussed in Section 6.
2. Install/configure NICs in the cluster nodes (three per node).
 - Public IP: As indicated by name.
 - Private interconnect: Use a dedicated high speed 1Gbps Ethernet connection between two cluster nodes using an Ethernet switch.
 - Server connection to NetApp storage: Connect one gigabit NIC to the gigabit switch that connects to the gigabit NIC on the NetApp storage.
3. Configure the network interfaces on each node.

- i. Configure the three network interfaces as follows:

```
# fredibm1 (Host 1)
en0 - ip: 172.17.4.185, netmask: 255.255.255.0
en1 - ip: 172.31.4.195, netmask: 255.255.255.0
en2 - ip: 192.168.10.1, netmask: 255.255.255.0

# fredibm2 (Host 2)
en0 - ip: 172.17.4.186, netmask: 255.255.255.0
en1 - ip: 172.31.4.196, netmask: 255.255.255.0
en2 - ip: 192.168.10.2, netmask: 255.255.255.0
```

Where:

Interface en0 is the public IP for each node.

Interface en1 on both cluster nodes is connected to the gigabit switch for storage I/O.

Interface en2 on both cluster nodes is connected back to back for cluster private interconnects.

- ii. Update the `/etc/hosts` file on the cluster nodes and add entries for public, private, and VIP addresses. In addition to the preconfigured public and private network, Oracle Database 11g requires additional IP addresses to be mapped to the public address as virtual IPs (VIPs). If a node fails when an application or user makes a connection using a VIP, Oracle Clusterware transfers 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 of `/etc/hosts` entries are as follows:

```
# Internet Address Hostname # Comments
172.17.4.85 fredibm1 fredibm1.iop.eng.netapp.com
172.17.4.86 fredibm2 fredibm2.iop.eng.netapp.com
192.168.10.1 fredibm1-priv
192.168.10.2 fredibm2-priv
172.17.4.185 fredibm1-vip fredibm1-vip.iop.eng.netapp.com
172.17.4.186 fredibm2-vip fredibm2-vip.iop.eng.netapp.com
172.31.4.189 fas3140c-sv111v fas3140c-sv111v.iop.eng.netapp.com
172.31.4.190 fas3140c-sv112v fas3140c-sv112v.iop.eng.netapp.com
```

4. Check the connectivity of each interface (interconnects, public IPs, and storage) using the `ping` command.
5. Create NFS mount points and mount the volumes with the following mount options on all cluster nodes. As a root user, update the `/etc/filesystems` file on all server nodes and add the following entries:

```
.....
..... df -k
```

```

.....

```

Filesystem	1024-blocks	Free	%Used	Iused	%Iused	Mounted on
/dev/hd4	131072	112692	15%	1956	8%	/
/dev/hd2	1376256	64048	96%	34468	66%	/usr
/dev/hd9var	327680	10700	97%	555	18%	/var
/dev/hd3	458752	123148	74%	1480	6%	/tmp
/dev/hd1	20054016	11426168	44%	30319	2%	/home
/proc	-	-	-	-	-	/proc
/dev/hd10opt	196608	14552	93%	3050	46%	/opt
172.31.4.189:/vol/vol1	301989888	255308608	16%	179	1%	/u01/app/oracle
172.31.4.189:/vol/vol2	20971520	18388548	13%	109	1%	/oralog
172.31.4.189:/vol/vol3	8388608	8075760	4%	101	1%	/storage
172.31.4.189:/vol/vol4	16777216	16563400	2%	101	1%	/oract1
172.31.4.190:/vol/vol5	247463936	207461936	19%	5778	1%	/oradata

Where:

- Data1 is the name of the NetApp storage system.
- /u01/app/oracle, /oralog, /storage, /oract1, and /oradata are the mount points on the cluster nodes.
- storage is just a separate mount point for CRS files. CRS files (cluster registry file and voting disk file) can reside in the same /storage volume but must be mounted with the `intr` mount option. The `intr` mount option is required in failure scenarios in which CRS has to evict a node.
- During the Oracle CRS install, make sure to indicate a path starting with /storage directory when prompted for the Oracle Cluster registry file (`ocr`) and voting disk (`css`) file location.

Note:

The `cio` mount option should not be used for `ORACLE_HOME` and `CRS_HOME` (the volume where CRS and Oracle binaries are placed).

Oracle installation will fail if you dynamically mount the NFS volumes without adding entries in the `/etc/filesystems` file. This is because Oracle Installer checks for entries and mount options in the `/etc/filesystems` file.

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

```

/u01/app/oracle

/oralog

/storage

/oract1

/oradata

```

7. Mount exported volumes on mount points created above on all the cluster nodes. Verify the mount options using the `mount` command on each node. After the NFS volumes are mounted, change the ownership of these mounted volumes to Oracle user and oinstall group.

8 INSTALLATION PROCEDURE

To install Oracle Database 11g RAC on AIX, refer to the Oracle installation guide specific to your Oracle version:

For Oracle Database 11g R2 see:

http://download.oracle.com/docs/cd/E11882_01/install.112/e10814/toc.htm

For Oracle Database 11g R1 see:

http://download.oracle.com/docs/cd/B28359_01/install.111/b28258/toc.htm

For more information on best practices and information on Clusterware, see:

http://www.oracle.com/technology/products/database/clusterware/pdf/rac_aix_system_stability.pdf

<http://www.oracle.com/technologies/grid/docs/clusterware-11g-whitepaper.pdf>

Note: There is a known issue (bug 8341679) with Oracle Database 11g and AIX 6.1. See metalink.oracle.com for more information.

9 APPENDIX: SAMPLE PROFILE FILE

Sample .profile file for the Oracle user:

```
export ORACLE_BASE=/u01/app/oracle;
export ORACLE_PRODUCT=$ORACLE_BASE/product;
export ORACLE_HOME=$ORACLE_PRODUCT/11.1.0/db_1;
export ORACLE_CRS=$ORACLE_PRODUCT/11.1.0/crs_1;
export ORACLE_SID=orcl;
export
LD_LIBRARY_PATH=$ORACLE_HOME/lib:$ORACLE_CRS/lib:$ORACLE_HOME/lib32:$LD_LIBRARY_PATH;
export LIBPATH=$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;
```

REVISION HISTORY

Author Name	Revision Date	Comments
Fred Schimscheimer	January 2007	Original document
Naveen Harsani	June 2010	Updated to include 11g contents

ARCHIVAL COPY
Contents may be out-of-date

NetApp provides no representations or warranties regarding the accuracy, reliability or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.