IBM

Technical report:

Oracle RAC10*g* with AIX and IBM System Storage N series

Best-practices installation

Document NS3542-0

October 10, 2007

Table of contents

Abstract

This technical report covers the installation of IBM System Storage N series in a network file system environment for IBM AIX on IBM System p server. The database is Oracle 10g Real Application Clusters Release 2. The focus of this paper is technical, and the reader should be comfortable with AIX and logical partition systems administration, Oracle 10g administration, network connectivity, and IBM N series system administration.

Introduction

This technical report covers the installation of IBM® System Storage™ N series in a network file system (NFS) environment for the IBM AIX® operating system on an IBM System p5™ 550Q server, model 9133-55A. The System p5 550Q Express platform is a new breed of 8-core powered by the latest IBM quadcore module (QCM) 1.65 GHz POWER5+™ processor cores, a single IBM System p5 550Q in multiple independent logical partitions (LPARs) running AIX Version 5.3 Technology Level 4 (5300-04) operating system. The database is Oracle10g Real Application Clusters Release 2 (hereafter referred to as OracleRAC10gR2) with Oracle Cluster Ready Services (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 OracleRAC10gR2 on IBM N series storage in the NFS environment.

Assumptions

We assume that readers are familiar with OracleRAC10gR2 and the operation of IBM N series storage systems. We also assume that readers are familiar with the operation of the IBM AIX environment and installation of Oracle 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 OracleRAC10*g*R2 certification environment specified by Oracle and IBM N series.
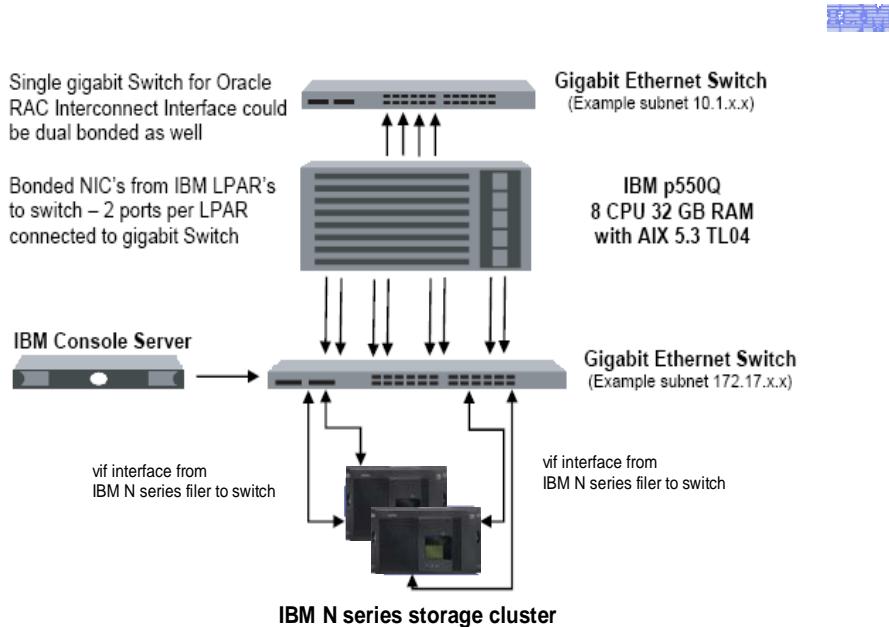
*Figure 1. OracleRAC10*g*R2 on IBM System p5 POWER5+ server with IBM N series storage.*

Figure 1 illustrates a typical configuration of OracleRAC10*g*R2 with IBM N series storage and IBM System p5 POWER5+ server running AIX Version 5.3. This is a scalable configuration and allows users to scale vertically or horizontally in terms of processor, memory, and storage.

As shown in the network diagram, we recommend that you dedicate a private network connection between the OracleRAC10*g*R2 servers and the N series storage storage system. This is accomplished using a dedicated gigabit network (with a gigabit switch) to the IBM N series storage cluster. A dedicated network connection is beneficial for the following reasons:

- In an OracleRAC10*g*R2 environment, it is important to eliminate any contentions and latencies.
- Providing a separate network ensures security.

The cluster interconnect is an essential part for Oracle database clusters. Along with cache fusion, the interconnect is used to monitor the heartbeat of the servers in the existing cluster group. This is a typical configuration that can be deployed in an enterprise environment.
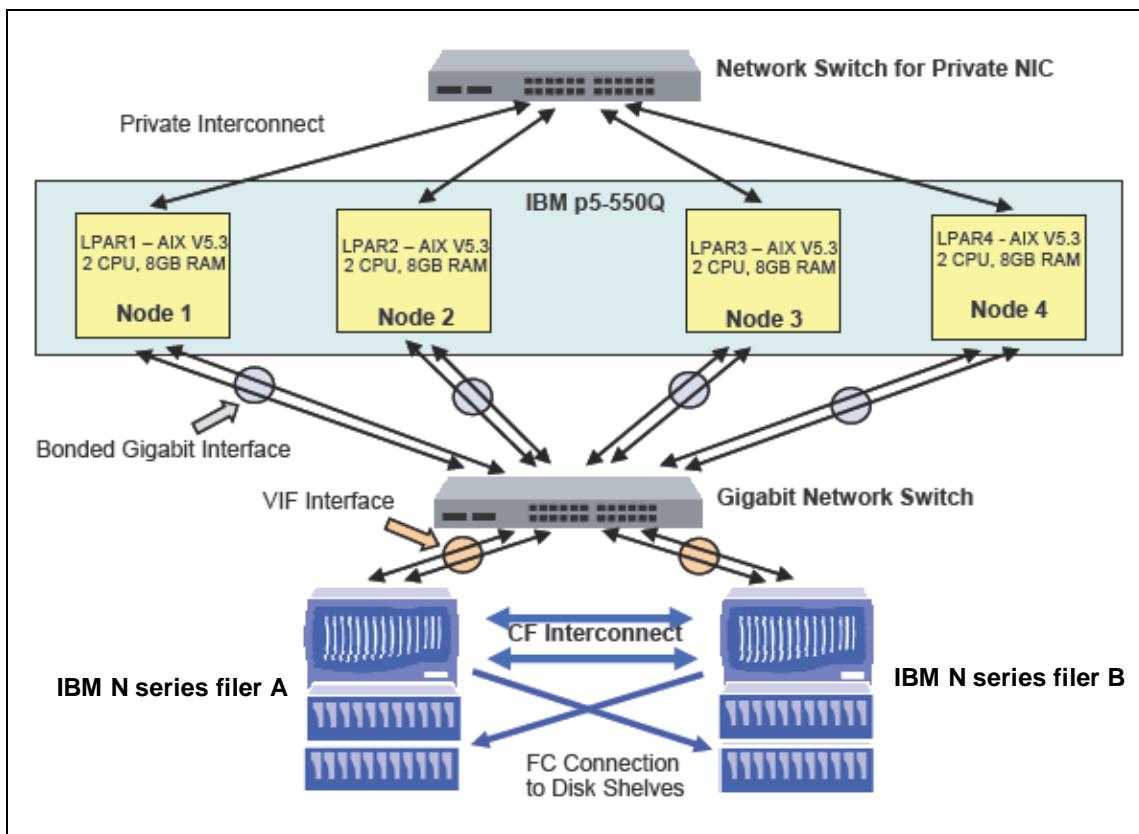
Figure 2. OracleRAC10gR2 configuration.

Requirements

Hardware used for tests

Cluster nodes

- One IBM System p5 550Q, Model 9133-55A, 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 the IBM N series)

Storage infrastructure

- One IBM System Storage N series 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 (5300-04)
- OracleRAC10gR2 (10.2.0.2) license.

Setup for IBM N series storage

1. Configure an IBM N series storage system running Data ONTAP 7.1 or later with NFS and SnapRestore® license keys.
2. Create and export volumes for storing Oracle database files on the storage:
   a. Create 9 volumes on the storage (Data1) in each LPAR as listed below:

```
/home1  Shared Oracle and CRS Home(Binaries)
/ora_data  Oracle datafiles
/ora_ctrl1 Active control files and CRS files
/ora_ctrl2 Standby control files and CRS files
/ora_log1 Active database logs, a copy of control file and archive logs
/ora_log2 Standby database logs, a copy of control file and archive logs
/vd1  Voting disk 1 with 100MB allocation
/vd2  Voting disk 2 with 100MB allocation
/vd3  Voting disk 3 with 100MB allocation
```

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

```
Data1> vol create vol1 14
```

   Note: Volume voldata was created with 14 disks and volumes vollog and volhome with eight disks each. You can create your volumes based on your workload needs.

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

```
/vol/vol1 -anon=0
/vol/vol2 -anon=0
/vol/vol3 -anon=0
/vol/vol4 -anon=0
/vol/vol5 -anon=0
/vol/vol6 -anon=0
/vol/vol7 -anon=0
/vol/vol8 -anon=0
/vol/vol9 -anon=0
```

   d. 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 has the ability to pool 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 OracleRAC10gR2 install, the following patches/file sets need to be applied on each IBM LPAR server. Some of these patches/file sets may already be applied to your system. Please verify if they already exist before applying them.

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

```
# lslpp -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 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)
- xlC.aix50.rte:7.0.0.4 or higher (see the following note)
- xlC.rte:7.0.0.1 or higher (see the following note)

The xlC file sets 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 V5.3 and V5.2. Therefore, Oracle10gR2 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.

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 limits 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, the "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 servers.

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 parameter settings over 64K)
```

These parameters can be set via 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 OracleRAC10*g*R2 functionality.

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 vmtune 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

  NAME                      CUR    DEF    BOOT   MIN    MAX    UNIT          TYPE
      DEPENDENCIES
  --------------------------------------------------------------------------------
  cpu_scale_memp            8      8      8      1      64                   B
  --------------------------------------------------------------------------------
  data_stagger_interval     161    161    161    0      4K-1   4KB pages     D
      lgpg_regions
  --------------------------------------------------------------------------------
  defps                     1      1      1      0      1      boolean       D
  --------------------------------------------------------------------------------
  force_relalias_lite       0      0      0      0      1      boolean       D
  --------------------------------------------------------------------------------
  framesets                 2      2      2      1      10                   B
  --------------------------------------------------------------------------------
  htabscale                 n/a    -1     -1     -4     0                    B
  --------------------------------------------------------------------------------
  kernel_heap_psize         4K     4K     4K     4K     16M    bytes         B
  --------------------------------------------------------------------------------
  large_page_heap_size      0      0      0      0      8E-1   bytes         B
      lgpg_regions
  --------------------------------------------------------------------------------
  lgpg_regions              0      0      0      0                           D
      lgpg_size
  --------------------------------------------------------------------------------
  lgpg_size                 0      0      0      0      16M    bytes         D
      lgpg_regions
  --------------------------------------------------------------------------------
  low_ps_handling           1      1      1      1      2                    D
  --------------------------------------------------------------------------------
  lru_file_repage           1      1      1      0      1      boolean       D
  --------------------------------------------------------------------------------
  lru_poll_interval         10     10     10     0      60000  milliseconds  D
  --------------------------------------------------------------------------------
  lrubucket                 128K   128K   128K   64K           4KB pages     D
  --------------------------------------------------------------------------------
  maxclient%                80     80     80     1      100    % memory      D
      maxperm%
      minperm%
  --------------------------------------------------------------------------------
  maxfree                   1088   1088   1088   8      200K   4KB pages     D
      minfree
      memory_frames
  --------------------------------------------------------------------------------
  maxperm                   1202K         1202K                             S
  --------------------------------------------------------------------------------
  maxperm%                  80     80     80     1      100    % memory      D
      minperm%
      maxclient%
  --------------------------------------------------------------------------------
  maxpin                    1266K         1266K                             S
  --------------------------------------------------------------------------------
  maxpin%                   80     80     80     1      99     % memory      D
      pinnable_frames
      memory_frames
  --------------------------------------------------------------------------------
  mbuf_heap_psize           64K    0      0      0      16M    bytes         B
  --------------------------------------------------------------------------------
  memory_affinity           1      1      1      0      1      boolean       B
  --------------------------------------------------------------------------------
  memory_frames             1568K         1568K                4KB pages     S
```

```
-------------------------------------------------------------------------------
memplace_data              2       2       2       1       2                     D
      memory_affinity
-------------------------------------------------------------------------------
memplace_mapped_file       2       2       2       1       2                     D
      memory_affinity
-------------------------------------------------------------------------------
memplace_shm_anonymous     2       2       2       1       2                     D
      memory_affinity
-------------------------------------------------------------------------------
memplace_shm_named         2       2       2       1       2                     D
      memory_affinity
-------------------------------------------------------------------------------
memplace_stack             2       2       2       1       2                     D
      memory_affinity
-------------------------------------------------------------------------------
memplace_text              2       2       2       1       2                     D
      memory_affinity
-------------------------------------------------------------------------------
memplace_unmapped_file     2       2       2       1       2                     D
      memory_affinity
-------------------------------------------------------------------------------
mempools                   2               2                                     d
      cpu_scale_memp
-------------------------------------------------------------------------------
minfree                    960     960     960     8       200K    4KB pages      D
      maxfree
      memory_frames
-------------------------------------------------------------------------------
minperm                    153859          153859                                S
-------------------------------------------------------------------------------
minperm%                   5       20      5       1       100     % memory       D
      maxperm%
      maxclient%
-------------------------------------------------------------------------------
nokilluid                  0       0       0       0       4G-1    uid            D
-------------------------------------------------------------------------------
npskill                    3968    3968    3968    1       496K-1 4KB pages       D
-------------------------------------------------------------------------------
npsrpgmax                  31K     31K     31K     0       496K-1 4KB pages       D
      npsrpgmin
-------------------------------------------------------------------------------
npsrpgmin                  23808   23808   23808   0       496K-1 4KB pages       D
      npsrpgmax
-------------------------------------------------------------------------------
npsscrubmax                31K     31K     31K     0       496K-1 4KB pages       D
      npsscrubmin
-------------------------------------------------------------------------------
npsscrubmin                23808   23808   23808   0       496K-1 4KB pages       D
      npsscrubmax
-------------------------------------------------------------------------------
npswarn                    15872   15872   15872   0       496K-1 4KB pages       D
-------------------------------------------------------------------------------
num_spec_dataseg           0       0       0       0                              B
-------------------------------------------------------------------------------
numpsblks                  496K            496K                    4KB blocks     S
-------------------------------------------------------------------------------
page_steal_method          0       0       0       0       1       boolean        B
-------------------------------------------------------------------------------
pagecoloring               n/a     0       0       0       1       boolean        B
-------------------------------------------------------------------------------
pinnable_frames            1443K           1443K                   4KB pages      S
```

```
-------------------------------------------------------------------------------
pta_balance_threshold      n/a    1     1     0     99      % pta segment     D
-------------------------------------------------------------------------------
relalias_percentage        0      0     0     0     32K-1                     D
-------------------------------------------------------------------------------
rpgclean                   0      0     0     0     1       boolean           D
-------------------------------------------------------------------------------
rpgcontrol                 2      2     2     0     3                         D
-------------------------------------------------------------------------------
scrub                      0      0     0     0     1       boolean           D
-------------------------------------------------------------------------------
scrubclean                 0      0     0     0     1       boolean           D
-------------------------------------------------------------------------------
soft_min_lgpgs_vmpool      0      0     0     0     90      %                 D
     lgpg_regions
-------------------------------------------------------------------------------
spec_dataseg_int           512    512   512   0                               B
-------------------------------------------------------------------------------
strict_maxclient           1      1     1     0     1       boolean           D
     strict_maxperm
-------------------------------------------------------------------------------
strict_maxperm             0      0     0     0     1       boolean           D
     strict_maxclient
-------------------------------------------------------------------------------
v_pinshm                   0      0     0     0     1       boolean           D
-------------------------------------------------------------------------------
vm_modlist_threshold       -1     -1    -1    -2    2G-1                      D
-------------------------------------------------------------------------------
vmm_fork_policy            1      1     1     0     1       boolean           D
-------------------------------------------------------------------------------
vmm_mpsize_support         1      1     1     0     1       boolean           B
-------------------------------------------------------------------------------

n/a means parameter not supported by the current platform or kernel

Parameter types:
    S = Static: cannot be changed
    D = Dynamic: can be freely changed
    B = Bosboot: can only be changed using bosboot and reboot
    R = Reboot: can only be changed during reboot
    C = Connect: changes are only effective for future socket connections
    M = Mount: changes are only effective for future mountings
    I = Incremental: can only be incremented
    d = deprecated: deprecated and cannot be changed
 ioo -L

NAME                       CUR    DEF   BOOT  MIN   MAX     UNIT             TYPE
     DEPENDENCIES
-------------------------------------------------------------------------------
j2_dynamicBufferPreallocation
                           16     16    16    0     256     16K slabs         D
-------------------------------------------------------------------------------
j2_inodeCacheSize          400    400   400   1     1000                      D
-------------------------------------------------------------------------------
j2_maxPageReadAhead        128    128   128   0     64K     4KB pages         D
-------------------------------------------------------------------------------
j2_maxRandomWrite          0      0     0     0     64K     4KB pages         D
-------------------------------------------------------------------------------
j2_maxUsableMaxTransfer    512    512   512   1     4K      pages             M
-------------------------------------------------------------------------------
j2_metadataCacheSize       400    400   400   1     1000                      D
-------------------------------------------------------------------------------
```

```
j2_minPageReadAhead        2     2     2     0     64K    4KB pages       D
-------------------------------------------------------------------------------
j2_nBufferPerPagerDevice   512   512   512   0     256K                   M
-------------------------------------------------------------------------------
j2_nPagesPerWriteBehindCluster
                           32    32    32    0     64K                    D
-------------------------------------------------------------------------------
j2_nRandomCluster          0     0     0     0     64K    16KB clusters   D
-------------------------------------------------------------------------------
j2_nonFatalCrashesSystem   0     0     0     0     1      boolean         D
-------------------------------------------------------------------------------
j2_syncModifiedMapped      1     1     1     0     1      boolean         D
-------------------------------------------------------------------------------
jfs_clread_enabled         0     0     0     0     1      boolean         D
-------------------------------------------------------------------------------
jfs_use_read_lock          1     1     1     0     1      boolean         D
-------------------------------------------------------------------------------
lvm_bufcnt                 9     9     9     1     64     128KB/buffer    D
-------------------------------------------------------------------------------
maxpgahead                 8     8     8     0     4K     4KB pages       D
     minpgahead
-------------------------------------------------------------------------------
maxrandwrt                 0     0     0     0     512K   4KB pages       D
-------------------------------------------------------------------------------
memory_frames              1568K       1568K               4KB pages     S
-------------------------------------------------------------------------------
minpgahead                 2     2     2     0     4K     4KB pages       D
     maxpgahead
-------------------------------------------------------------------------------
numclust                   1     1     1     0     2G-1   16KB/cluster    D
-------------------------------------------------------------------------------
numfsbufs                  196   196   196   1     2G-1                   M
-------------------------------------------------------------------------------
pd_npages                  64K   64K   64K   1     512K   4KB pages       D
-------------------------------------------------------------------------------
pgahd_scale_thresh         0     0     0     0     1254K  4KB pages       D
-------------------------------------------------------------------------------
pv_min_pbuf                512   512   512   512   2G-1                   D
-------------------------------------------------------------------------------
sync_release_ilock         0     0     0     0     1      boolean         D
-------------------------------------------------------------------------------

n/a means parameter not supported by the current platform or kernel

Parameter types:
    S = Static: cannot be changed
    D = Dynamic: can be freely changed
    B = Bosboot: can only be changed using bosboot and reboot
    R = Reboot: can only be changed during reboot
    C = Connect: changes are only effective for future socket connections
    M = Mount: changes are only effective for future mountings
    I = Incremental: can only be incremented
    d = deprecated: deprecated and cannot be changed
```

Because OAST, a complete database stress test kit, uses a lot of processes, kernel tuning was done on nfiles and nprocesses using smitty. The default number of processes is 128 and changed it to 100,000.

```
SW_dist_intr   false              Enable SW distribution of interrupts         True
autorestart    true               Automatically REBOOT system after a crash    True
boottype       disk               N/A                                          False
capacity_inc   0.01               Processor capacity increment                 False
capped         true               Partition is capped                          False
conslogin      enable             System Console Login                         False
cpuguard       enable             CPU Guard                                    True
dedicated      false              Partition is dedicated                       False
ent_capacity   1.00               Entitled processor capacity                  False
frequency      528000000          System Bus Frequency                         False
fullcore       false              Enable full CORE dump                        True
fwversion      IBM,SF240_174      Firmware version and revision levels         False
id_to_partition 0X80000EEC5F300002 Partition ID                                False
id_to_system   0X80000EEC5F300000 System ID                                    False
iostat         false              Continuously maintain DISK I/O history       True
keylock        normal             State of system keylock at boot time         False
max_capacity   4.00               Maximum potential processor capacity         False
max_logname    9                  Maximum login name length at boot time       True
maxbuf         20                 Maximum number of pages in block I/O BUFFER CACHE True
maxmbuf        0                  Maximum Kbytes of real memory allowed for MBUFS True
maxpout        0                  HIGH water mark for pending write I/Os per file True
maxuproc       100000             Maximum number of PROCESSES allowed per user  True
min_capacity   1.00               Minimum potential processor capacity         False
minpout        0                  LOW water mark for pending write I/Os per file True
modelname      IBM,9133-55A       Machine name                                 False
ncargs         6                  ARG/ENV list size in 4K byte blocks          True
pre430core     false              Use pre-430 style CORE dump                  True
pre520tune     disable            Pre-520 tuning compatibility mode            True
realmem        6422528            Amount of usable physical memory in Kbytes   False
rtasversion    1                  Open Firmware RTAS version                   False
sed_config     select             Stack Execution Disable (SED) Mode           True
systemid       IBM,02109BE9F      Hardware system identifier                   False
variable_weight 0                 Variable processor capacity weight           False
```

Preinstall setup tasks (cluster nodes)

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

1. Ensure that all four LPARs of the IBM System p5 550Q, model 9133-55A server have the latest recommended patches/file sets and operating-system settings as discussed in the section entitled, "Patches, environment, and operating-system settings."
2. Install/configure NICs in the cluster nodes (three per node).
   a. Public internet protocol (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 the 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.
   d. Configure the three network interfaces as indicated below.

```
# fredibm1 (Host 1)
en0 – ip: 172.17.4.185, netmask: 255.255.255.255
en3 - ip: 172.31.4.195, netmask: 255.255.255.255
en1 - ip: 192.168.10.1, netmask: 255.255.255.255

# fredibm2 (Host 2)
en0 – ip: 172.17.4.186, netmask: 255.255.255.255
en3 - ip: 172.31.4.196, netmask: 255.255.255.255
en1 - ip: 192.168.10.2, netmask: 255.255.255.255

# fredibm3 (Host 3)
en0 – ip: 172.17.4.187, netmask: 255.255.255.255
en3 - ip: 172.31.4.197, netmask: 255.255.255.255
en1 - ip: 192.168.10.3, netmask: 255.255.255.255

# fredibm4 (Host 4)
en0 – ip: 172.17.4.188, netmask: 255.255.255.255
en3 - ip: 172.31.4.198, netmask: 255.255.255.255
en1 - ip: 192.168.10.4, netmask: 255.255.255.255
```

Where:

- Interface en0 is the public IP for each node.
- Interface en3 on both cluster nodes is connected to the gigabit switch for storage I/O.
- Interface en1 on both cluster nodes is connected back to back for cluster private interconnects.

e. Update the /etc/hosts file on the cluster nodes and add entries for public, private, and virtual IP (VIP) addresses. Note that in addition to the preconfigured public and private network, Oracle10*g* 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

172.17.4.85  fredibm1 fredibm1.iop.eng.ibmn.com
172.17.4.86  fredibm2 fredibm2.iop.eng.ibmn.com
172.17.4.87  fredibm3 fredibm3.iop.eng.ibmn.com
172.17.4.88  fredibm4 fredibm4.iop.eng.ibmn.com
192.168.10.1 fredibm1p
192.168.10.2 fredibm2p
192.168.10.3 fredibm3p
192.168.10.4 fredibm4p
172.17.4.185 fredibm1v fredibm1v.iop.eng.ibmn.com
172.17.4.186 fredibm2v fredibm2v.iop.eng.ibmn.com
172.17.4.187 fredibm3v fredibm3v.iop.eng.ibmn.com
172.17.4.188 fredibm4v fredibm4v.iop.eng.ibmn.com
172.31.4.189 fas980c-svl11v fas980c-svl11v.iop.eng.ibmn.com
172.31.4.190 fas980c-svl12v fas980c-svl12v.iop.eng.ibmn.com
```

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

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:

```
.....
.....    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% /ora_data
172.31.4.189:/vol/vol2    20971520  18388548   13%      109     1% /ora_log2
172.31.4.189:/vol/vol3     8388608   8075760    4%      101     1% /vd1
172.31.4.189:/vol/vol4     8388608   8075772    4%      101     1% /vd3
172.31.4.189:/vol/vol5    16777216  16563400    2%      101     1% /ora_ctrl1
172.31.4.190:/vol/vol6   100663296  98115308    3%      109     1% /ora_log1
172.31.4.190:/vol/vol7    33554432  33338840    1%      101     1% /ora_ctrl2
172.31.4.190:/vol/vol8     8388608   8075772    4%      101     1% /vd2
172.31.4.190:/vol/vol9   247463936    167012  100%     5778     1% /home1
.....
.....    mount
.....
  node       mounted          mounted over    vfs      date          options
-------- ---------------  ---------------  ------ ------------ ---------------
        /dev/hd4          /                jfs2   Dec 12 11:23 rw,log=/dev/hd8
        /dev/hd2          /usr             jfs2   Dec 12 11:23 rw,log=/dev/hd8
        /dev/hd9var       /var             jfs2   Dec 12 11:23 rw,log=/dev/hd8
        /dev/hd3          /tmp             jfs2   Dec 12 11:23 rw,log=/dev/hd8
        /dev/hd1          /home            jfs2   Dec 12 11:23 rw,log=/dev/hd8
        /proc             /proc            procfs Dec 12 11:23 rw
        /dev/hd10opt      /opt             jfs2   Dec 12 11:23 rw,log=/dev/hd8
172.31.4.189 /vol/vol1    /ora_data        nfs3   Dec 12 11:23
cio,rw,bg,hard,nointr,proto=tcp,vers=3,noac,rsize=32768,wsize=32768,timeo=600
172.31.4.189 /vol/vol2    /ora_log2        nfs3   Dec 12 11:23
cio,rw,bg,hard,nointr,proto=tcp,vers=3,noac,rsize=32768,wsize=32768,timeo=600
172.31.4.189 /vol/vol3    /vd1             nfs3   Dec 12 11:23
cio,rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600
172.31.4.189 /vol/vol4    /vd3             nfs3   Dec 12 11:23
cio,rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600
172.31.4.189 /vol/vol5    /ora_ctrl1       nfs3   Dec 12 11:23
rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600
172.31.4.190 /vol/vol6    /ora_log1        nfs3   Dec 12 11:23
cio,rw,bg,hard,nointr,proto=tcp,vers=3,noac,rsize=32768,wsize=32768,timeo=600
172.31.4.190 /vol/vol7    /ora_ctrl2       nfs3   Dec 12 11:23
rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600
172.31.4.190 /vol/vol8    /vd2             nfs3   Dec 12 11:23
cio,rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600
172.31.4.190 /vol/vol9    /home1           nfs3   Dec 12 11:23
rw,bg,hard,nointr,proto=tcp,vers=3,noac,rsize=32768,wsize=32768,timeo=600
.....
.....    lsfs -l
.....
Name            Nodename   Mount Pt              VFS      Size   Options   Auto Accounting
/dev/hd4        --         /                    jfs2    262144     --      yes  no
/dev/hd1        --         /home                jfs2  40108032     --      yes  no
/dev/hd2        --         /usr                 jfs2   2752512     --      yes  no
```

```
/dev/hd9var      --             /var            jfs2   655360      --     yes no
/dev/hd3         --             /tmp            jfs2   917504      --     yes no
/proc            --             /proc           procfs    --       --     yes no
/dev/hd10opt     --             /opt            jfs2   393216      --     yes no
/vol/vol1      172.31.4.189 /ora_data          nfs3      --
cio,rw,bg,hard,nointr,proto=tcp,vers=3,noac,rsize=32768,wsize=32768,timeo=600 yes  no
/vol/vol2      172.31.4.189 /ora_log2          nfs3      --
cio,rw,bg,hard,nointr,proto=tcp,vers=3,noac,rsize=32768,wsize=32768,timeo=600 yes  no
/vol/vol3      172.31.4.189 /vd1               nfs3      --
cio,rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600      yes  no
/vol/vol4      172.31.4.189 /vd3               nfs3      --
cio,rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600      yes  no
/vol/vol5      172.31.4.189 /ora_ctrl1         nfs3      --
rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600          yes  no
/vol/vol6      172.31.4.190 /ora_log1          nfs3      --
cio,rw,bg,hard,nointr,proto=tcp,vers=3,noac,rsize=32768,wsize=32768,timeo=600 yes  no
/vol/vol7      172.31.4.190 /ora_ctrl2         nfs3      --
rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600          yes  no
/vol/vol8      172.31.4.190 /vd2               nfs3      --
cio,rw,bg,hard,intr,proto=tcp,vers=3,rsize=32768,wsize=32768,timeo=600      yes  no
/vol/vol9      172.31.4.190 /home1             nfs3      --
rw,bg,hard,nointr,proto=tcp,vers=3,noac,rsize=32768,wsize=32768,timeo=600   yes  no
```

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 just 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. The "intr" mount 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: The "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 the /etc/filesystems file.

6. Create the following mount points on all cluster nodes:
```
#mkdir /ora_data
#mkdir /home1
#mkdir /ora_ctrl1
#mkdir /ora_ctrl2
#mkdir /ora_log1
#mkdir /ora_log2
#mkdir /vd1
#mkdir /vd2
#mkdir /vd3
```
7. Mount the exported volumes on the mount points created above on all the cluster nodes. It is always a good idea to verify mount options by mount command on each node. After the NFS volumes are mounted, change the ownership of these mounted volumes to Oracle user.

Installation procedure

For an OracleRAC10*g* R2 install on AIX, please refer to Oracle install documentation.

Appendix

Sample `.profile` **file for the Oracle user:**

```
export ORACLE_BASE=/home1/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:$LD_LIBRARY_PATH;
export LIBPATH=$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;
```

Trademarks and special notices

Each environment has its own specific set of requirements and no guarantees can be given that the results presented in this report will work as expected on other platforms. This paper should assist in the research and troubleshooting that may be required in a particular case and serve as a checklist of items of which to be aware.