

Oracle 9i & 10gR1 on IBM AIX5L: Tips & Considerations



October 2009
IBM Oracle International Competency Center
San Mateo, California

Contents

Introduction	3
Operating system Considerations	4
Technology-Level / Maintenance-Level / filesets	4
Kernel	5
Storage considerations	7
Raw Logical Volumes	7
Filesystems.....	8
Filesystem IO Tuning.....	9
Alerts issues & workarounds.....	12
Release 9i:.....	12
Release 10g:	14
References.....	17
DISCLAIMER	17
TRADEMARKS.....	18

INTRODUCTION

The objective of this document is to provide considerations and recommendations to practitioners involved with the installation of Oracle®'s RDBMS product on IBM's AIX 5L operating system. Though not intended to function as a step-by-step implementation guide, this paper, when used in conjunction with published materials from both Oracle and IBM, will be an asset to you as you plan and implement a successful Oracle RDBMS solution for your customer.

This paper contains a range of recommendations available at time of publication for implementing single-instance Oracle RDBMS solutions. Note: a detailed discussion of Oracle Real Application Clusters, (RAC), is not within the scope of this document. However, with the inclusion of content relating to release 10g added to these pages, some RAC related commentary is appropriate. Information specific to Oracle 10g will be included in a dedicated section of this document. Other areas addressed in these pages are base operating system set-up, and storage considerations. Although these topics are presented separately, it is important to recognize that they are tightly integrated in preparing an RDBMS environment, and further that they are independent of the Oracle database release deployed. The content has been provided by IBM technical support specialists experienced with Oracle products in AIX environments, and published technical references.

Please leverage <http://metalink.oracle.com>, or <http://otn.oracle.com/support/metalink/content.html> for certified Oracle solution information, as well as the latest installation and product release notes. Oracle's Administrative and Product Release documentation will also provide an initial starting point for base operating system preparation and tuning recommendations. The "References" section of this paper provides specific identification of these resources.

This final version of the document corresponds to the final patch set releases of Oracle 9iR2 and 10gR1. For information about newer releases of Oracle's Server on AIX refer to these documents:

- *"Oracle Database 10gR2 and Oracle RAC 10gR2 on IBM AIX: Tips and Considerations"*, <http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101089>
- *"Oracle Database 11g and Oracle RAC 11g on IBM AIX: Tips and Considerations"*, <http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101176>

Your feedback is important for us. We want our technical papers to be as helpful as possible. Please send your comments about this document to the IBM/Oracle International Competency Center (IOICC). The IOICC's email address is:

ibmoracl@us.ibm.com

OPERATING SYSTEM CONSIDERATIONS

Preparation of the AIX operating system is critical for a high-performing and stable Oracle RDBMS environment. Discussion of Maintenance Level, base operating system, (bos), file sets, and C/C++ compiler requirements follows.

TECHNOLOGY-LEVEL / MAINTENANCE-LEVEL / FILESETS

Regardless of the AIX5L release, (5.2-5.3, 5.1 was desupported on 16-JUN-2006) present on your host system, apply the latest Technology Level (TL) (`#oslevel -r`) to your base operating system. In addition, we encourage you to apply the most recently released service pack (`#oslevel -s`). If you have not applied ML01 (Maintenance Level terminology applies through AIX 5300-03) to your AIX 5.3 environment, please apply the following mandatory PFTs for the APARs

IY58143 - Required for optimal system stability and performance
IY59386 – “libdepend.mk files are all empty”
IY58396 - rename block_transfer param
IY60930 - Unable to remote routes

APARs IY66513 and IY70159 should be applied to AIX 5.3 for any application where “LDR_CNTRL” environment variable is explicitly set.

APAR IY68989 is mandatory for all Oracle customers running on AIX V5.3 ML01 or ML02. This fix address a critical problem introduced in ML01. The symptoms can vary from an Oracle application hang to unexplained corruption of Oracle jar files or Oracle Application Forms.

APAR IY68874 addresses the issue “An application that is running with mandatory large page data (LDR_CNTRL=LARGE_PAGE_DATA=M) may core-dump on the AIX 5.3 64-bit kernel in a LPAR environment.

APAR IY70031 fixes the issue “CORRUPTION FROM SIMULTANEOUS CIO WRITES WITH O_DSYNC ON JFS2”.

Note: All customers at AIX 5.3 TL05 (5300-05) must install the IBM AIX PTF for APAR IY89080, and obtain the fix for Oracle Bug 5496862. Please see Metalink Doc id’s 390656.1, and 5496862.8 for further information.

Please note that IBM has made enhancements to the service strategy for AIX. For more information on this topic, please see “[The AIX 5L Operating System: Service Strategy Details and Best Practices](#)”. This publication is available from the IBM “Service and support best practices” website:

<http://www14.software.ibm.com/webapp/set2/sas/f/best/home.html>.

Discussion of TL’s, SP’s, Concluding Service Packs (CSP), and interim fixes (IF) is included.

In addition to the most recent Technology Level / Maintenance Level, we encourage you to apply the latest Service Pack (SP) for AIX operating system level. You can obtain TL’s, SP’s, and ML packages from:

<http://www-03.ibm.com/servers/eserver/support/unixservers/aixfixes.html>. The Latest TL available in AIX 5L version 5.3 when this document is updated is TL10.

The following steps will help you navigate the site to obtain your Technology Level package or Service Pack respectively:

- Choose the appropriate AIX release link from “Fix Bundles” from the “Quick links for AIX fixes” page
- On the resulting AIX “Fix Bundles for AIX operating system” page, select your desired level for your AIX operating system TL or SP click the appropriate filename. On the resulting page, you can review the package information and pursue download
- Agree to the terms and conditions to proceed from the “Download” tab. Provide your current AIX level (`oslevel -r`) via the drop down box on the “Packaging options” page and continue

- Review your package listing on the "Package results" page and choose your download option – Java applet or ftp
- Follow the prompts to acquire the package via your selected download mechanism

Please review Metalink Document ID# **Note:282036.1**

"Minimum software versions and patches required to Support Oracle Products on IBM pSeries". This dynamic posting reflects minimum software versions to enable the latest Oracle product offerings on AIX5L. This document will identify the base operating system (bos) filesets required for Oracle, as well as making recommendations for specific fixes, by ML/TL. We recommend you leverage this document in the building of your AIX Operating System image for your Oracle product.

Though the above listed Metalink note will provide you with packages required for your product install, please review your release appropriate documentation for additional environment preparation information. Please navigate to and review the "Operating System Requirements" and "System Requirements" sections, respectively, of the following documents:

Oracle 9iR2: <http://otn.oracle.com/documentation/oracle9i.html>

1. "Oracle9i Installation Guide Release 2 for UNIX Systems: AIX-Based Systems, Compaq Tru64 UNIX, HP 9000 Series HP-UX, Linux Intel, and Sun Solaris", (part# A96167-01)
2. "Oracle9i Release Notes Release 2 (9.2.0.1.0) for AIX-Based 5L Systems (64-bit)", (part# A97605-06)

Oracle 10gR1: <http://www.oracle.com/technology/documentation/database10g.html>

1. "Oracle Database Installation Guide for UNIX Systems", (part# B10811-05)
2. "Oracle Database Release Notes 10g Release 1 (10.1.0.2.0) for AIX-Based Systems", (part# B13611-08)
3. "Oracle Real Application Clusters Installation and Configuration Guide" (part# B10766-08)

Oracle 10gR2: <http://www.oracle.com/technology/documentation/database10gr2.html>

1. "Oracle Database Installation Guide 10g Release 2 (10.2) for AIX 5L Based Systems (64-Bit)" (part# B19075-02) http://download-east.oracle.com/docs/cd/B19306_01/install.102/b19075.pdf
2. "Oracle Database Release Notes 10g Release 2 (10.2) for AIX 5L Based Systems (64-Bit)", (part# B19074-07) http://download-east.oracle.com/docs/cd/B19306_01/relnotes.102/b19074.pdf
3. "Oracle Database Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10g Release 2 (10.2) for AIX" (part# B14201-04) http://download-east.oracle.com/docs/cd/B19306_01/install.102/b14201/toc.htm

In addition to the previously mentioned AIX maintenance strategy and Metalink publication #282036.1, please review the appropriate Oracle publication(s) for operating system requirements.

The IBM C and C++ compiler run-time environment is shipped on AIX installation media and installed along with base AIX. Install the latest runtime maintenance when installing AIX. When the XL C/C++ Enterprise Edition for AIX compiler is installed, install the latest compiler maintenance. The latest compiler maintenance can be downloaded by pointing a browser at <http://www-306.ibm.com/software/awdtools/xlcpp/features/aix/> and clicking on the "Latest XL C/C++ Updates (PTFs)" link. On the "Latest updates for supported IBM C and C++ compilers" page that appears, follow the appropriate link under AIX. If no compiler is installed, follow the "runtime only" link. Follow instructions provided to download and install the maintenance.

KERNEL

The AIX 5L operating system, supports both 32-bit and 64-bit applications. The only requirement to run 64-bit applications on an AIX system is that the hardware be 64-bit and that the 64-bit Application Environment is enabled - see below. All Oracle products, with the exception of Oracle 10gR2 RAC on AIX 5.3, operate in either a 32-bit or 64-bit kernel environment. The 64-bit kernel is required when the system, or LPAR, has more than 96 GB of real memory. Please review other 3rd party software product restrictions for 32-bit or 64-bit kernel. Check with the appropriate vendors for further details.

To determine if your hardware is 64-bit, enter:

```
#bootinfo -y
```

A "64" will be displayed if the hardware is in fact 64-bit. Enable 64-bit application mode using the 'smitty load64bit' fast path. Please note that the "bos.64bit" file set (Base Operating System 64 bit Runtime) must be installed before 64-bit application mode can be enabled. Enabling 64-bit application mode allows 64-bit applications, such as Oracle 9i to run. The 64-bit mode setting has no effect on the functionality of 32-bit applications. Please re-apply Maintenance or Technology levels to update the bos.64bit file set as required if you are adding the bos.64bit file set subsequent to any ML / TL application on your system.

With the AIX 5L operating system, you have the flexibility of booting to a 64-bit kernel. Entering the following command in an AIX 5L environment will indicate which kernel is loaded:

```
#bootinfo -K
```

Please note that some Oracle components need Oracle's pw-syscall kernel extension. Some older Oracle products ship with a kernel extension that supports only 32-bit kernels and an Oracle patch is required to obtain a 64-bit version of the kernel extension for use on the 64-bit kernel. For instance, 8i, 11i, and 9iAS release 1.0.2.2 require a patch to run on 64-bit kernel, whereas more recent products, (such as Oracle9i release 2, 9.2, or 9iAS 9.0.2), do not use an AIX kernel extension, and therefore do not require patching. Refer to Oracle's Metalink for patch numbers providing 64-bit kernel support for specific Oracle products. For example, Oracle patch #2896876 provides a modified 'rootpre.sh' to load the pw-syscall64 kernel extensions for Oracle 8.1.7.

STORAGE CONSIDERATIONS

The following section discusses technologies that when properly tuned and configured, can provide you with a reliable, stable I/O subsystem. Regardless of whether you implement your database on filesystems or raw volumes, incorporate I/O performance into your database at design time. By paying careful attention toward the design of the database's datafile placement, you can reduce the potential for I/O bottlenecks and the potential for a major rework on your database to address future performance issues.

By eliminating the I/O buffering overhead associated with filesystems, best performance under most conditions is achieved using AIX raw logical volumes. However, with advancements made in filesystems, the performance disparity is narrowing. The conveniences afforded by the ease of manageability and the continuing performance gains, have made filesystems popular in many environments.

Additional discussion of the trade-offs between raw logical volumes and filesystems on AIX are discussed in *"Oracle9i Release 2 (9.2.0.1.0) Administrator's Reference for UNIX Systems: AIX, Based Systems, Compaq Tru64 UNIX, HP 9000 Series HP-UX, Linux Intel, and Sun Solaris"*, May 2002, Part No. A97297-01. This publication is available at http://download-east.oracle.com/docs/pdf/A97297_01.pdf

However, no discussion of direct I/O (DIO) or concurrent I/O (CIO) with AIX takes place. Please note that Direct I/O (DIO) support has been available in AIX V5.1 since March 26, 2002, when PTF for APAR IY28122 "ADD MOUNT OPTION TO FORCE DIRECT I/O" was delivered. AIX V5.2 featured DIO support when it first became available on October 18, 2002. Please refer to the following comments on filesystems for information on these features.

For reference on these topics in context of release 10g, please refer to the *"Oracle® Database Administrator's Reference 10g Release 1 (10.1) for UNIX Systems: AIX-Based Systems, Apple Mac OS X, hp HP-UX, hp Tru64 UNIX, Linux, and Solaris Operating System Part No. B10812-06"*, available at http://download-west.oracle.com/docs/pdf/B10812_06.pdf.

AIX 5L fixes on multiple issues with MPIO PCM for the IBM Storage servers DS3/4/5K is available in PTF IZ29856, IZ28763, IZ31161 and IZ26373

RAW LOGICAL VOLUMES

When using AIX raw logical volumes for your Oracle datafiles, please use the zero offset feature, which is supported with either Big or Scalable Volume Groups. The capability to use zero offset logical volumes within Big Volume Groups (mkgv -B) was introduced in AIX 5.1 ML04 (IY36656), and AIX 5.2 ML01 (IY38077). As described in Metalink Note 261460.1 (bug 2620053), when Big Volume Groups are used, the Raw Logical Volumes must be created with the "mklv -T O" option. When Scalable Volume Groups (mkgv -S) are used, the mkgv "-T O" option is not required. In conjunction with AIX support for the zero offset logical volume, you will need to verify that your Oracle code is at the appropriate level to recognize the new device type. Oracle 9.2.0.3+ includes the fix. If you are down-level from this release, please apply the correct Oracle patch for your 9i version for bug #2620053.

The following APARs, included in AIX 5.3 TL06, are important to consider when preparing your AIX environment for raw logical volume support:

- IY94343 MKLV -TO ON BIG VOLUME GROUPS FAILS TO PUT SOME LV INFORMATION
- IY91832 CPLV DOES NOT COPY 1ST 1024 BYTES FROM SVG LOGICAL VOLUME

AIX 5L version 5.3 TL10 is available with latest fixes.

Note: For Oracle RAC environments, HACMP is required for all 9i implementations, as well as 10g RAC deployments where raw logical volumes are the chosen data storage solution. Please review Oracle's Metalink Certify "General Notes For RAC for Unix On IBM AIX based Systems (RAC only):" page for information on HACMP supported versions with Oracle. Additional information on support for HACMP can be found on Metalink: "Status of Certification of Oracle Clusterware with HACMP 5.3 & 5.4 Doc ID: Note:404474.1"

FILESYSTEMS

When choosing a filesystem solution, JFS2 with Concurrent I/O, (CIO), is the best performing filesystem for many Oracle workloads. In some situations however, there may be benefits of file system caching (e.g. sequential read ahead), so Direct I/O (DIO), or CIO may not always provide the best performance. CIO, available only on JFS2, requiring AIX 5.2 ML01 or later, allows filesystems to have performance characteristics reasonably close to that of raw devices.

As noted in the Oracle9i Release 2 (9.2.0.1.0) Administrator's Reference for UNIX Systems, Oracle9i takes full advantage of asynchronous I/O (AIO) provided by AIX, resulting in faster database access. The Oracle DISK_ASYNC_IO parameter should be left at its default value, "TRUE" (10g). However, in some situations, disabling this parameter may assist in problem resolution. The Administrator's Reference provides tuning recommendations for AIX AIO. When following those recommendations, change/verify the "STATE to be configured at system restart" field on the "Change / Show Characteristics of Asynchronous I/O" menu (accessed via the 'smitty chgaio') from defined to available.

If using Oracle8i or Oracle9i release 1 on JFS or JFS2, please apply fsync bugfix, Oracle patch 2024097. This fix is incorporated within subsequent Oracle 9i releases.

JFS

If using JFS instead of JFS2, CIO is not available but DIO is – use it. Do not allocate a JFS with the large file enabled (bf) attribute. The big file attribute increases the minimum DIO transfer size (diocapbuf.dio_min returned by the finfo system call) from 4K to 128K, forcing Oracle to read and write a minimum of 128K bytes to exploit DIO. Please note that 2 GB is the largest file size applications (including Oracle) can use in a JFS without the "bf" attribute - large file enablement. Do not allocate a compressed JFS, which defeats DIO. Respecting the JFS 2GB maximum file size avoids potential file level inode locking performance issues in JFS or JFS2 (without CIO), but a downside of small (2GB) files in very large database (VLDB) environments is the time required to open/close files during instance startup/shutdown

Please note that the "[Guidelines for Tuning File Systems](#)" article in the AIX v5.2 "[Performance Management Guide](#)" says, "JFS is optimized for 32-bit kernel, and will run on 32-bit and 64-bit hardware.

JFS2

JFS2 is supported on the 32-bit AIX kernel, though the AIX v5.3 "[Performance Management Guide](#)" states that "Enhanced JFS is the default file system for 64-bit kernel environments. Due to address space limitations of the 32-bit kernel, Enhanced JFS is not recommended for use in 32-bit kernel environments". As mentioned in the "Operating System Considerations: Maintenance-Level / Filesets" portion of this document, we strongly encourage you to apply the latest available maintenance levels for your AIX 5.2 or 5.3 environment. While encouraging you to apply the latest available Maintenance Level for your AIX release, we identify ML02 for 5.2 as a minimum starting point for JFS2 deployment. Please note that AIX must be at 5.2 ML01 or later for CIO. Because the CIO mount option was not in base AIX V5.2, it may not be covered in the initial release of AIX 5.2 documentation.

The CIO mount option should only be used for filesystems containing data which is intended to be accessed in concurrent mode, such as Oracle datafiles, online redo logs and control files, and should not contain libraries and executables, such as the filesystem containing the \$ORACLE_HOME or \$ORA_CRS_HOME. Please apply APAR #/Y56024 "CIO Write Returns Incorrect Length" to address ORA-600 errors that had occurred when tempfiles were located on CIO mounted JFS2 filesystems. Within an Oracle 10g environment, note that CIO/DIO is enabled at the file open level, as opposed to the filesystem level as in 9i. Therefore, the explicit mount option at the filesystem level is not required in 10g.

GPFS

GPFS is a fully supported cluster filesystem option for your Oracle 9i release 2 or 10g R1 RAC database.

GPFS 3.1 is certified with AIX5L 5.3 and has extended support for 10g R1 RAC and 9i RAC..

Review Metalink Note: 302806.1 "[IBM General Parallel File System \(GPFS\) and Oracle RAC on AIX 5L and IBM eServer pSeries](#)" for additional information on GPFS. This document covers all topics relevant to GPFS and Oracle RAC including GPFS support/certification, features for RAC, HACMP requirements, and implementation / tuning information. This Metalink note also provides installation examples to assist you with deployment.

For additional cluster filesystem information, we encourage you to view Metalink Doc ID: Note:183408.1 "[Raw Devices and Cluster Filesystems With Real Application Clusters](#)". This document provides an overview of the cluster storage options available for RAC including, raw devices, ASM, and cluster filesystems.

FILESYSTEM IO TUNING

Below are several basic recommendations for initial tuning of your filesystems for improved performance. Please refer to the [Oracle9i Release 2 \(9.2.0.1.0\) Administrator's Reference for UNIX Systems](#) or "[Oracle® Database Administrator's Reference 10g Release 1 \(10.1\) for UNIX Systems: AIX-Based Systems, Apple Mac OS X, hp HP-UX, hp Tru64 UNIX, Linux, and Solaris Operating System Part No. B10812-06](#)" for additional information on these and other recommended filesystem parameter settings for your environment.

When evaluating DIO/CIO, it is important to consider the impact this may have on other utilities and applications that may be accessing the filesystems housing your Oracle datafiles and logs. Many filesystem utilities, e.g. 'cp', 'dd', 'cpio', are not DIO/CIO aware. Where `filesystemio_options=setall`, these utilities will fail if issuing a concurrent open without the `O_CIO` or `O_DIRECT` flag. You can view the online documentation for additional comments:

<http://publib.boulder.ibm.com/infocenter/pseries/v5r3/index.jsp?topic=/com.ibm.aix.basetechref/doc/basetrf1/open.htm>

Another relevant consideration point for filesystem utilities within a DIO/CIO environment is that these utilities typically rely on a small default transfer size. When filesystem caching is used, filesystem read ahead and/or write coalescing effectively increases the actual I/O transfer size. This mitigates the potential performance issues with a small I/O transfer size. However, when DIO or CIO is used, the filesystem can no longer use read ahead or write coalesce. If the default utility I/O transfer size isn't overridden (e.g. `dd bs=512k`), throughput performance will be significantly degraded.

If DIO/CIO is used, make sure to set `DB_FILE_MULTIBLOCK_READ_COUNT` for Oracle. Setting `DB_FILE_MULTIBLOCK_READ_COUNT` will enable Oracle to perform read-ahead on sequential scans. A large value for the `DB_FILE_MULTIBLOCK_READ_COUNT` initialization parameter usually yields better I/O throughput. On AIX, this parameter ranges from 1 to 512, but using a value higher than 16 usually does not provide additional performance gains. Set this parameter so that its value when multiplied by the value of the `DB_BLOCK_SIZE` parameter produces a number that is larger than the LVM stripe size. Such a setting causes more disks to be used.

On CIO-mounted filesystems that contain datafiles, the database block size needs to be multiple of `agblksize`. This is not something that customers need to worry about except if the database block size is 2KB, in which case you need to make sure `agblksize` is 2KB, 1KB or 512 bytes.

Place Oracle online redo logs and control files in their own filesystem(s) with `agblksize` set to 512. If placing the Oracle Database's logs on a JFS2 filesystem, the optimal configuration is to create the file system with option `agblksize=512` and to mount it with option '`cio`'. This delivers logging performance similar to the performance of a raw device.

Where neither CIO nor DIO are used as filesystem mount options, assign `maxpgahead` (JFS) or `j2_maxPageReadAhead` (JFS2) = $[(DB_FILE_MULTIBLOCK_READ_COUNT * DB_BLOCK_SIZE) / 4KB]$ as

a starting point. Please refer to the AIX [Performance Management Guide](#) for further information on sequential page read ahead tuning.

Set **filesystemio_options** parameter in your init*.ora file to a value appropriate for your 9i or 10g database workload:

Oracle 9i

filesystemio_options	JFS(default mount)	JFS(dio mount)	JFS2(default mount)	JFS2(cio mount)
asynch	asynchronous, cached IO	asynchronous, DIO	asynchronous, cached IO	asynchronous, CIO
setall	asynchronous, cached IO	asynchronous, DIO	asynchronous, chased IO	asynchronous, CIO
directio*	synchronous, cached IO	synchronous, DIO	synchronous, cached IO	synchronous, CIO
none*	synchronous, cached IO	synchronous, DIO	synchronous, cached IO	synchronous, CIO

Oracle 10g

filesystemio_options	JFS(default mount)	JFS(dio mount)	JFS2(default mount)	JFS2(cio mount)
asynch	asynchronous, cached IO	asynchronous, DIO	asynchronous, cached IO	asynchronous, CIO
setall	asynchronous, DIO	asynchronous, DIO	asynchronous, CIO	asynchronous, CIO
directio*	synchronous, DIO	synchronous, DIO	synchronous, CIO	synchronous, CIO
none*	synchronous, cached IO	synchronous, DIO	synchronous, cached IO	synchronous, CIO

* - not recommended, always use asynch I/O

Metalink note # 272520.1 "Direct I/O or Concurrent I/O on AIX 5L" is a useful reference for values of this parameter. Please review the "Alerts Issues and Workarounds" portion of this document, as there are issues associated with this parameter. Note: Please be advised that if you leverage the AIX filemon command (http://publib16.boulder.ibm.com/doc_link/en_US/a_doc_lib/cmds/aixcmds2/filemon.htm), that AIO calls to aio_read/aio_write & [lio_listio](#), requests are not included in the logical filemon report.

Tune *minservers*/*maxservers*/*maxreqs* to accommodate the amount of I/O:

minservers: Oracle Corporation recommends an initial value equal to the number of CPUs on the system or 10, whichever is lower.

maxservers: Starting with AIX 5L version 5.2, this parameter counts the maximum number of AIO servers per CPU, whereas on previous versions of AIX it used to be a system-wide value. If using GPFS, set maxservers to worker1 threads divided by the number of CPUs, this will be the optimal setting and increasing maxservers will not lead to additional I/O performance. If using JFS/JFS2, set the initial value to (10 * number of logical disks / number of CPUs) and monitor the actual number of aioservers started during a typical workload using the pstat or ps commands. If the actual number of active aioservers is equal to the maxservers, then increase the maxservers value.

maxreqs: Set the initial value to (4 * number of logical disks * queue depth). You can determine the queue depth by running the following command: "\$ lsattr -E -l hdisknn"

AIX default values are typically inadequate for Oracle I/O requirements. A shortage of configured system resources to handle queued I/O requests typically manifests itself in DBWR trace message “DBWR TRACE WARNING: LIO_LISTIO RETURNED EAGAIN”. There are several sources for more information on asynchronous I/O tuning:

- 'Warning: lio_listio returned EAGAIN' On ASM Doc ID: Note:430697.1
- AIX: DBWR trace warning: LIO_LISTIO RETURNED EAGAIN Doc ID: Note:265491.1
- AIX v5.3 Performance Management Guide
http://publib.boulder.ibm.com/infocenter/pseries/v5r3/index.jsp?topic=/com.ibm.aix.prftungd/doc/prftungd/async_disk_io_perf_tuning.htm

We also encourage initial configuration values for AIX virtual memory manager, vmm. The recommendation here centers on AIX page replacement and the impact of `lru_file_repage`. Please review these topics in the AIX 5.3 Performance Management Guide (<http://publib.boulder.ibm.com/infocenter/pseries/v5r3/topic/com.ibm.aix.prftungd/doc/prftungd/prftungd.pdf>), and “Oracle Architecture and Tuning on AIX” (<http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100883>) for information on these and other AIX tuning topics.

Please note that when deriving values for the above parameters that “CPU” refers to the logical CPU’s on the system or LPAR in question – not physically allocated CPU resources.

AIXTHREAD_SCOPE. Oracle recommends the usage of system-wide contention, which maps Oracle threads to kernel threads statically. This is critical in an Oracle RAC environment, and provides the benefit of reduced memory allocation at the Oracle process level in AIX version 5.2 or greater. Set (export) “AIXTHREAD_SCOPE=S” as an environment variable for your Oracle instance owner.

These are some recommended starting points for your environment. Continue to monitor and tune your environment according to the requirements of your workload. Typically, there is little impact from setting these parameters too high, whereas too low a setting can result in a negative I/O performance impact.

For **GPFS**, please leverage the example in Metalink Doc id# 302806.1. This posting provides recommendations on LUN configuration, filesystem blocksize, AIO configuration, inodes, as well as implementation examples. Some notable initial settings and consideration points:

1. prefetchThreads & worker1Threads. You can leverage the example provided in the Metalink note for initial settings of these parameters – 75 and 475 respectively. Please review the posting for complete information on these variables.
2. GPFS blocksize. A larger blocksize is recommended for optimal performance – typically 512KB. The GPFS blocksize should not be set to DB blocksize.
3. Inode capacity. If your GPFS is intended to support a shared Oracle Home, and or CRS Oracle Home, the default number of maximum inodes will likely be inadequate.

We strongly encourage you to review 302806.1 directly for more detail on the above items and others for your GPFS deployment. Additional information on GPFS with Oracle is available online: <http://publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp?topic=/com.ibm.cluster.gpfs.doc/gpfsbooks.html>

ALERTS ISSUES & WORKAROUNDS

This portion of the document identifies several known bugs with Oracle 9i Release 2 and 10g Release 1 patch sets. Please leverage the following resources for more detail: Oracle Doc id's #**189908.1** "[ALERT: Oracle9i Release 2 \(9.2\) Support Status and Alerts](#)", or Oracle Doc id #**263719.1** "[ALERT: Oracle 10g Release 1 \(10.1\) Support Status and Alerts](#)". You can pursue these documents directly by submitting the appropriate Doc id number on the search bar. These documents will provide you with links to all subsequent patch sets released for the respective base Oracle RDBMS release. An alternative approach to locate these resources:

- Login to Oracle Metalink(<http://metalink.oracle.com>)
- Click on the 'Advanced' search hypertext (RH side of search box), to go to Oracle's Advanced Site Search page
- In the 'using all the words' box, type 'Support Status and Alerts' (including the quotation marks)
- Select your product from the 'Product' drop-down i.e. "Oracle Server – Enterprise Edition"
- Select your platform from the Platform drop-down i.e. "AIX5L Based Systems (64-bit)"
- Click the 'Submit' button

In the resulting hit list, find the Support Status and Alerts entry for the version of Oracle you are planning to implement. Within the "Patch Sets" section of the base release alert document, select the "List of fixes added in" link to view a listing of all fixes within a specific patch set. You can identify port specific fixes by the "P" suffix affixed to the bug number. Please note, that this site does not provide visibility into bugs identified subsequent to the most recently released patch set.

When viewing a Support Status and Alerts entry, be sure to note the Platform Specific Addenda section for AIX near the bottom. Peruse the hit list for other relevant hits.

RELEASE 9i:

Patch set 9.2.0.8:

This marks the terminal patch set release for 9i. Please refer to the Metalink Notes identified below for detail fix information as well as known issues and alerts. Two notable issues affecting the AIX environment are noted below:

Known issues and alerts affecting 9.2.0.8 Note: 388281.1

5659663P AIX: libskgpxu.a missing from 9.2.0.8 Patch Set. From the bug detail on Metalink: "libskgpxu.a is missing from the 9.2.0.8 Patch Set on AIX 5L. This shared library is required for RAC installations. The patch for this bug# can be downloaded from Metalink and should be applied after 9.2.0.8 has been installed if using RAC".

5507032P AIX/HPUX: "opatch" not installed with 9.2.0.8 on AIX / HPUX

OPatch is not installed with the 9.2.0.8 patchset on these platforms:

AIX5L
HPUX (PA-RISC)
HPUX (Itanium)

Workaround:

Download and install the latest version of OPatch from Metalink (patch 2617419)

Patch set 9.2.0.7:

9.2.0.7 Patch Set - Availability and Known Issues Note: 308895.1

Patch set 9.2.0.6:

Bug #4079525 - Applying any patch after 9206 patch set will not succeed on Sun, AIX, HP

Workaround: See published workaround in Metalink Doc ID: Note:292946.1

Bug #4074633 – Higher CPU usage in RAC background processes

Bug #4081980 - ORA-06508: PL/SQL error running dbca after applying the 9206 patch set

Workaround: Apply the patch for bug #4081980.

Bug #3910149 - RDBMS Server: Patch KGLHDDEP SGA LEAK IN RAC.

This issue is believed to affect all RDBMS releases < 10.2. It is fixed in 9.2.0.7, 10.1.0.5, and 10.2.0.1. There is a patch available for this bug.

Patch set 9.2.0.5:

Bug #3733456 - After Upgrade from 9.2.0.3 To 9.2.0.5 On Aix5l Database Direct I/O Not Working

Bug 3917158 - ORA-29740 and a false instance eviction can occur

Fixed in 9.2.0.6. Please see Metalink Doc ID: Note:3917158.8 for additional information.

Bug 3580447 - AIX: HSODBC does not work in 9.2.0.5 on AIX.

Regression introduced in 9.2.0.5. Fixed in 9.2.0.6

Bug #3917158 - ORA-29740 and a false instance eviction can occur <Note:3917158.8>

This issue is fixed in 9.2.0.6. A patch has been provided for 9.2.0.4 & 9.2.0.5. Please see related bugs # bug 4206159, and 3648788 for **10g** related fixes.

Bug #3393861 – “GSD does not start on AIX 5L”

Patch set 9.2.0.4:

Bug #3415744 - GSD JRE VERSION 1.1.8 NOT COMPATIBLE WITH AIX DYNAMIC LPAR: The IBM LPAR utility refuses to dynamically add resources to the LPAR because of the version of JRE it uses (1.1.8).

Workaround: Submit a TAR for a fix/workaround referencing this bug number.

Bug #3264088 - DIRECT LOAD MAY BE SLOW DUE TO REPETED FILE OPEN/CLOSE

Note: Previous bug filings #2509662 & #3145019 are resolved with this patch.

Bug #3224502 - IDENTIFYING A JFS2 FILE SYSTEM TYPE AS A MMFS:

Oracle incorrectly recognizes JFS2 filesystems as GPFS and therefore forces Direct I/O for JFS2 filesystems. This may or may not be desirable in your environment. There is no fix available from Oracle yet. Please see Doc ID: Note:281253.1 for additional information.

Workaround: In the absence of a fix, the problem can be circumvented by adding the following line to /etc/vfs:

```
mmfs 8 /usr/lpp/mmfs/bin/mmfsmnthelp none
```

Please note that the descriptive text in the Oracle Bug report 3224502 on Metalink does not detail the behavior, nor the workaround as it is described here.

Patch set 9.2.0.3:

There is a known filesystem issue introduced in 9.2.0.3. You will experience issues with filesystem performance because Oracle9i will use Synchronous I/O to access filesystems, even if Asynchronous I/O is turned on. This bug may cause performance degradation whether or not Concurrent I/O is turned on.

To find out if you are impacted by the above issue, (if you're on 9.2.0.3 you can be 99% sure that you are), simply truss the process that is doing I/O. If you don't see it using `lio_listio()` to submit I/O, it's using Sync I/O instead of Async I/O. Bugs filed for this issue:

Bug #2510406 - AIX 5L: FILESYSTEMIO_OPTIONS is set to NONE with a default on AIX:

Verify by running "show parameter filesystemio_options" command. To avoid poor performance, would seem prudent to set "filesystemio_options=async" in init.ora except for:

Bug #2603538 - AIX 5L: FILESYSTEMIO_OPTIONS parameter is not read from the init.ora file:

Workaround: This bug is fixed in patch set 9.2.0.5. Please submit a TAR through Oracle Metalink to request a fix.

Bug #3048661 - CREATE BITMAP INDEXES TAKES LONGER IN 9203 WHEN PGA => 2GB

In some customer environments, manual PGA memory management was found to be better performing than automatic PGA memory management during the creation of very large bitmap indexes.

Workaround: This bug is fixed in 9.2.0.5 and 10g base release

General Issues with 9i on AIX:

Doc ID# 259983.1: Memory Consumption on AIX

Please leverage this note for several tips on reducing high levels of memory consumption running Oracle on AIX.

Doc ID# 293715.1: After upgrading to AIX 5.2 ML-4 Oracle JAVAVM fails permanently

This issue affects Oracle versions 9.2.0.1 – 9.2.0.6. Customers experiencing an 'ORA-29516' should apply AIX e-fix IY63366.

Doc ID#:261460.1: AIX Logical Volume use can Corrupt Oracle Data (Bug 2620053)

This issue affects:

- All 8.1.7 patch sets ;
- All 9.0.1 patch sets ;
- 9.2.0.1 ;
- 9.2.0.2 .

The issue is fixed in 9.2.0.3 onwards but requires an OS level patch and reconfiguration of the database in order for the fix to take effect.

RELEASE 10g:

As mentioned above, please reference Oracle Doc id #263719.1 "*ALERT: Oracle 10g Release 1 (10.1) Support Status and Alerts*" for issues associated to your particular 10g release. The following listing identifies patches that we recommend you apply as standalone , if you have not reached the patch set level that contains the fix, within the '<>' following the bug description where applicable. Subsequent sections are ordered by patch set release.

Bug #3648788 - CORE: Patch CRS DEAMON SHUTDOWN THE NODE WHEN THE SYSTEM TIME IS ADJUSTED FORWARD

Bug #3626461 - CRS HUNG AFTER REBOOTING ON MASTER NODE. <Fixed in 10.1.0.4>

Bug #3953420 - MLR: ONS ROLLUP ON TOP OF 10.1.0.3.

Bug #3455036 - crsd may not start on reboot. <Fixed in 10.1.0.4>

Bug #3809246 - RAC-CRITICAL: TC3A: CRSD MEMORY LEAK. <Fixed in 10.1.0.4>

Bug #3476889 - EXPDP fails with OERI[ktcsetscn-1] with large SCN value. <Fixed in 10.1.0.4>

Bug #3942568 - NODE CRASHING WHEN HEARTBEAT CABLE IS PULLED ON OTHER (MASTER) NODE. <Fixed in 10.1.0.4>

Bug #3555800 - 10G CLIENT CAN'T CONNECT VIA ASO TO 10G DB ON DIFFERENT ENDIAN ORDER PLATFORM –("Cannot connect 10g client to server on different platform with AES256"). <Fixed in 10.1.0.4>

Patch set 10.1.0.5

Bug # 5021340 - INTERNAL SQL GENERATED WHEN USING JMSCORRELATIONID IN A SELECTOR IS UNSHAREABLE

Bug #6708377 - DST CHANGES FOR AUSTRALIA / EGYPT/VENEZUELA/BRAZIL/ARGENTINA – NEED OJVM FIX

Bug #4472437 - PSE FOR BASE BUG 4472437 ON TOP OF 10.1.0.5 FOR HPUX 64BIT (59)

Bug #5675477 - MERGE LABEL REQUEST ON TOP OF 10.1.0.5 FOR BUGS 4966417 4359124

Bug #4505133 - 10.1.0.5 PATCH SET FOR ORACLE DATABASE SERVER

Patch set 10.1.0.4:

Bug #4437469 – RACGVIP NOT FUNCTIONAL WITH ETHERNET VIRUTALIZATION

At the time of publication for this document, this was an unpublished bug. Please submit an Oracle TAR if you encounter this issue. We believe it to be a regression introduced in 10.1.0.4 patch set, and that a workaround exists.

Bug #3779593 - CRSD may crash on a surviving node after communication error with a failed node. This is fixed in 10.0.1.4.

Patch set 10.1.0.3:

Bug #4217746 - LD: 0706-006 CANNOT FIND OR OPEN LIBRARY FILE: -L ONS

Description: Oracle 10.1.0.3.0 on AIX 5L. Single node server with a non-Rac installation.

Symptoms: ORACLE_HOME/bin> relink network_

Error: ld: 0706-006 Cannot find or open library file: -l ons ld:open(): A file or directory in the path name does not exist. make: 1254-004 The error code from the last command is 255. .

ORACLE_HOME/bin> lsnrctl 0 bytes. Cause: The ORACLE_HOME/bin/libons.a file is missing. Fix: Copied the file from an in-house server and emailed to customer. We recommend you search for the bug and if a fix is not posted, file a TAR referencing this bug number.

As 10g continues to mature, we expect to see more activity in the patch space to address issues as they're encountered in the field. Please use Oracle Metalink's advanced search procedure below to assist you in examining the available patches for this offering:

PATCHES (example, 10.1.0.3):

<http://metalink.oracle.com>

->Patches

->Advanced Search

Product or Product Family -> select 'flashlight' / spreadsheet icon

-> select 'Search By Hierarchy'

-> expand 'Database and Tools'

-> select 'Oracle Database Family'

-> select checkbox 'Include Products belonging to selected Product Families'

-> Release = 'Oracle 10.1.0.3' (flashlight icon for lookup) 10.1.0...

-> Platform or Language = 'AIX5L Based Systems (64-bit)'

-> Patch Type = 'Any'

Oracle 10g database for AIX 5L, (releases 5.2 and 5.3 only), provides new features that are unique from previous releases. Additional features such as Automatic Storage Management (ASM), and Cluster Ready Services (CRS) increase the configuration flexibility of single instance and RAC deployments. As mentioned previously, we strongly recommend that you review Oracle Metalink (<http://metalink.oracle.com>) note #282036.1 for the latest AIX operating system levels and APAR recommendations to support Oracle products on IBM system p servers. Also, please view Oracle's Certification pages to explore supported configuration options.

ASM

Automatic Storage Management (ASM) is a technology newly introduced in 10g that provides for filesystem and volume management facilities within the Oracle database engine. ASM is designed to function in both RAC and single instance databases.

For technical information on ASM we recommend you refer to: "Oracle Database10g Automatic Storage Management Technical Best Practices", which is accessible via URL from Metalink Doc ID: Note:265633.1 Subject: ASM Technical Best Practices.

CRS

Oracle Cluster Ready Services comprises the clusterware for RAC deployments of 10g. With CRS included with 10g RAC, HACMP is no longer a pre-requisite for RAC. However, where customers desire raw logical volumes for their database files, HACMP will be required to provide concurrent raw logical volume support.

On section 2, page 40 of "[Oracle Real Application Clusters Installation and Configuration Guide](#)" 10g release 1 (10.1) for AIX – Based Systems Part No. B10766-08", you will find a discussion on leveraging raw logical volumes to support the CRS registry and voting devices. HACMP's Enhanced Concurrent Logical Volume Management (ECLVM) is a pre-requisite for pursuing this storage configuration option for CRS devices, as well as for raw logical devices for Oracle datafiles. The documentation indicates that the activation of the raw devices for CRS will be performed by CRS startup script - init.crs. At this time however, init.crs is not capable of varying on the concurrent volume group and we recommend that you pursue physical 'hdisk' devices for cluster registry and voting purposes.

When installing CRS on a node where you have HACMP file sets installed, but no cluster topography defined, you may experience failure of CRS services to start on that node. This scenario would only occur if you elect not to use HACMP as part of your 10g RAC deployments, but the file sets reside on your system. Simply remove the HACMP software, and your 'root.sh' should succeed if all other pre-requisites for CRS have been met.

If you experience issues with your CRS install, please observe Oracle's recommended clean-up procedure as described in Metalink note # 239998.1 entitled "[10g RAC: How to Clean Up After a Failed CRS Install](#)" before attempting a subsequent CRS installation.

REFERENCES

- [Oracle9i Administrator's Reference for UNIX Systems: AIX-Based Systems #A97297-01](http://otn.oracle.com/documentation/oracle9i.html) Available at: <http://otn.oracle.com/documentation/oracle9i.html>
- [Oracle9i Release Notes Release 2 \(9.2.0.1.0\) for AIX-Based 5L Systems \(64-bit\) #A97605-06](http://otn.oracle.com/documentation/oracle9i.html) Available at: <http://otn.oracle.com/documentation/oracle9i.html>
- [Oracle on AIX FAQ documents: #'s 231901.1, 210073.1, 223521.1, and 223522.1](http://metalink.oracle.com) Available at: <http://metalink.oracle.com>
- [IBM eServer pSeries Support: http://www-1.ibm.com/servers/eserver/support/pseries/](http://www-1.ibm.com/servers/eserver/support/pseries/)
- [Configuring the IBM Enterprise Storage Server for Oracle OLTP Applications](http://www.storage.ibm.com/disk/ess/pdf/oltpapp.pdf) Available at: <http://www.storage.ibm.com/disk/ess/pdf/oltpapp.pdf>
- [AIX v5.3 Performance Management Guide](http://publib.boulder.ibm.com/infocenter/pseries/v5r3/topic/com.ibm.aix.prftungd/doc/prftungd/prftungd.pdf) Available at: <http://publib.boulder.ibm.com/infocenter/pseries/v5r3/topic/com.ibm.aix.prftungd/doc/prftungd/prftungd.pdf>
- Information about Concurrent I/O is available in the "[Improving Database Performance With AIX Concurrent I/O, A case study with the Oracle9i database on AIX 5L version 5.2](http://www-1.ibm.com/servers/aix/whitepapers/db_perf_aix.pdf)" whitepaper at: http://www-1.ibm.com/servers/aix/whitepapers/db_perf_aix.pdf
- [Oracle Database10g Automatic Storage Management Technical Best Practices, Nitin Vengurlekar, Oracle Corporation, January 2004](http://www.virtual.com/whitepapers/Oracle_10g_Auto_Storage_Management_wp.pdf#search='oracle%20ASM%20%20best%20practices') Available at: http://www.virtual.com/whitepapers/Oracle_10g_Auto_Storage_Management_wp.pdf#search='oracle%20ASM%20%20best%20practices'
- [Oracle Database Administrator's Reference for UNIX Systems #B10812-06](http://download-west.oracle.com/docs/pdf/B10812_06.pdf) Available at: http://download-west.oracle.com/docs/pdf/B10812_06.pdf
- [Oracle Architecture and Tuning on AIX](http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP100883) Available at: <http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP100883>
- [Oracle Database 10gR2 and RAC 10gR2 on IBM AIX 5L: Tips and Considerations,](http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP101089) <http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP101089>
- [Oracle Database 11g and RAC 11g on IBM AIX 5L: Tips and Considerations,](http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP101176) <http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP101176>

DISCLAIMER

This document reflects the IBM / Oracle International Competency Center's understanding about Oracle products running on IBM hardware servers and storage. It was produced and reviewed by the members of the IBM organization called the IBM / Oracle International Competency Center and others.

This document is presented "As-Is" and IBM does not assume responsibility for the statements expressed herein. It reflects the experiences of the IBM / Oracle International Competency Center. If you have questions about the contents of this document, please direct them to the IBM / Oracle International Competency Center (ibmoracl@us.ibm.com).

TRADEMARKS

All brand names are the trademarks of their respective owners.



© IBM Corporation 2004
IBM Corporation
Marketing Communications
Systems and Technology Group
Route 100
Somers, New York 10589

Produced in the United States of America

October 2004
All Rights Reserved

This document was developed for products and/or services offered in the United States. IBM may not offer the products, features, or services discussed in this document in other countries. The information may be subject to change without notice. Consult your local IBM business contact for information on the products, features and services available in your area.

All statements regarding IBM future directions and intent are subject to change or withdrawal without notice and represent goals and objectives only.

IBM, the IBM logo, the e-business logo, ~, AIX, AIX 5L, HACMP, POWER and pSeries are trademarks or registered trademarks of International Business Machines Corporation in the United States or other countries or both. A full list of U.S. trademarks owned by IBM may be found at:
<http://www.ibm.com/legal/copytrade.shtml>.

UNIX is a registered trademark of The Open Group in the United States, other countries, or both.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

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

Information concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of the non-IBM products should be addressed with the suppliers.

The IBM home page on the Internet can be found at <http://www.ibm.com>.

The pSeries home page on the Internet can be found at
<http://www.ibm.com/servers/eserver/pseries>.