

Oracle 9i on IBM AIX5L Tips & Considerations

Version 5.6

May, 2004



**IBM / Oracle International Competency Center
San Mateo, California**

IBM

Contents

Introduction	3
Operating system Considerations	4
Maintenance-Level / filesets	4
Kernel	4
Storage considerations	6
Filesystems.....	6
Filesystem IO Tuning.....	7
Alerts issues & workarounds.....	9
References.....	11
DISCLAIMER	11
TRADEMARKS.....	11

INTRODUCTION

The objective of this document is to provide 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. Please note that a discussion of Oracle9i Real Application Clusters, (RAC), is not within the scope of this document. Areas addressed include: 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. 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. The 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.

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), filesets, and C/C++ compiler requirements follows.

MAINTENANCE-LEVEL / FILESETS

Regardless of the AIX5L release present on your host system, apply the latest **Maintenance Level** to your base operating system. Maintenance Levels at time of publication for AIX 5.1 and 5.2 are ML06 and ML03, respectively.

Following the application of the most recent Maintenance Level, please apply critical AIX fixes to your environment. You can identify and download critical fixes by pointing your browser at

<http://www-912.ibm.com/eserver/support/fixes/fcgui.jsp>

- Choose "pSeries, RS/6000" from the drop-down on the "Fix Central" page
- Choose "AIX OS, Java, compilers" from the "Product or fix type" drop down
- Choose "Critical fixes" from the "Option" drop down
- Choose your operating system level from the "OS level" drop down

Note: These selection drop down boxes appear in succession, and may not appear if a choice has not been made.

- Click on the "Continue" button
- On the resulting "Critical fixes" page, select the latest available date from the drop down and click "Go"
- Download and install all fixes, "Get all"

Please navigate to <http://otn.oracle.com/documentation/oracle9i.html> and review the "Operating System Requirements" and "System Requirements" sections, respectively, of the following documents:

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)*

Please review your environment and apply any filesets & PTFs mentioned within these documents.

Additionally, the IBM C and C++ compiler run-time environment is shipped on AIX installation media and installed along with base AIX. If the compiler run-time environment will be used (by, for example, Oracle), install the latest run-time maintenance when installing AIX. When a compiler is installed, install the latest compiler maintenance. The latest compiler maintenance can be downloaded by pointing a browser at <http://www.ibm.com/software/awdtools/vacpp/support/> and clicking on the "Latest Updates (PTFs)" link. On the "Supported C and C++ compilers: links to latest updates (PTFs)" 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 – Oracle9i for example. The only requirement to run 64-bit applications on an AIX system is that the hardware be 64-bit. All Oracle products, including 64-bit 9i, 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 fileset (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 apply maintenance levels to update the bos.64bit fileset as required if you are adding the bos.64bit fileset subsequent to maintenance level application(s) 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 iAS 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" (dated May 2002), is available at <http://otn.oracle.com/documentation/oracle9i.html>.

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.

Note: Please include a thorough examination of the "Alerts Issues & Workarounds" section of this document when assessing any of the following topics.

FILESYSTEMS

When choosing a filesystem solution, JFS2 with Concurrent I/O, (CIO), is the best performing filesystem for most 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". 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 the "STATE to be configured at system restart" field on the "Change / Show Characteristics of Asynchronous I/O" menu (accessed via the 'smitty chgaio' fastpath) 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. However, the AIX v5.2 "[Performance Management Guide](#)" states that enhanced JFS (i.e. JFS2) is optimized to run on the 64-bit kernel, allowing it to take advantage of 64-bit functionality. 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 5.1 or AIX 5.2 environment. At time of publication, we identify ML-05 and ML-02 for AIX 5.1 and 5.2 respectively, to provide a baseline starting point for jfs2 deployment, and to address JFS2 issues previously identified within these pages.

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 initial release of the 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, log files and control files, and should not contain libraries and executables, such as the filesystem containing the ORACLE_HOME. Please apply APAR #IY56024, "CIO Write Returns Incorrect Length", to address ORA-600 errors that had occurred when tempfiles were located on CIO mounted JFS2 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](#) for additional information on these and other recommended filesystem parameter settings for your environment

If DIO/CIO is used, make sure to set *DB_FILE_MULTIBLOCK_READ_COUNT* for Oracle to perform the read-ahead. 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 gain. 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.

Where standard filesystem mount option is used, please 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.

Place Oracle logs on their own filesystem 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.

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 *worker1threads* 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 $(10 * \text{number of logical disks} / \text{number of CPUs})$ and monitor the actual number of *aio* servers started during a typical workload using the *pstat* or *ps* commands. If the actual number of active *aio* servers is equal to the *maxservers*, then increase the *maxservers* value.

maxreqs: Set the initial value to $(4 * \text{number of logical disks} * \text{queue depth})$. You can determine the queue depth (typically 3), by running the following command: "\$ lsattr -E -l hdiskxx"

ALERTS ISSUES & WORKAROUNDS

This portion of the document identifies several known bugs with Oracle 9i Release 2. Though organized by 9.2.0.x release, please review all bugs identified in this section for relevance. You can identify platform specific bugs on Oracle Doc id #189908.1 "[ALERT: Oracle9i Release 2 \(9.2\) Support Status and Alerts](#)".

Navigate to this document in the following manner:

- Login to Oracle Metalink(<http://metalink.oracle.com>)
- Click on the advanced search button to go to Oracle's Advanced Site Search page
- In the keywords box, type 'Support Status and Alerts' (including the quotation marks)
- Select your product from the product Search Attribute drop-down i.e. "Oracle Server – Enterprise Edition"
- Click the Search button

In the resulting hit list, find the Support Status and Alerts entry for the version of Oracle you are planning to implement. The above referenced alert document addresses the base 9.2 version and all subsequent patch levels released for 9.2. Select the "List of fixes added in 9.2.0.x" Oracle note to view a listing of all fixes. You can identify platform 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.

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 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 patchset 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: For optimal filesystem performance, apply the 9.2.0.4 patchset
Please review the specifics of the bug, and refer to Metalink Doc Id#223299.1 for further information on `pga_aggregate_target` parameter settings.

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. In the absence of a fix, the problem can be circumvented by adding the line: "mmfs 8 usr/lpp/mmfs/bin/mmfsmnhelp none" to /etc/vfs.

Please note that the descriptive text in the Oracle Bug report 3224502 on Metalink does not detail the behavior, nor the fix provided as its stated here.

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:

Please see the published workaround for this bug on Oracle's Metalink site.

Bug #3264088 - DIRECT LOAD PERFORMANCE PROBLEM DUE TO REPEATED FILE OPEN
AND Note: Previous bug filings #2509662 & #3145019 are resolved with this patch.

Doc id# 259983.1: Memory Consumption on AIX

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

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, 206569.1, 223521.1, and 223522.1 Available at: <http://metalink.oracle.com>
- IBM eServer pSeries Support: <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>
- Information about Direct I/O is available in the "Tuning Direct I/O" article in the "Performance Management Guide" in the AIX V5.1 web-based documentation at: http://publibn.boulder.ibm.com/doc_link/en_US/a_doc_lib/aixbman/prftungd/2365c813.htm
- 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" whitepaper at: http://www-1.ibm.com/servers/aix/whitepapers/db_perf_aix.pdf

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.