Oracle Database 11*g* R2
Oracle RAC 11*g* R2 on IBM AIX

Tips and Considerations

Ravisankar Shanmugam
IBM Oracle International Competency Center
October 2011

Table of contents

Change history **3**
Abstract **4**
Introduction **4**
Terminology **4**
IBM AIX **5**
AIX support 6
C and C++ compilers 7
Oracle Database 11*g* Release 2 ..... **7**
Current certifications 8
Latest patch set 9
Virtualization for Oracle Database 11*g* Release 2 10
Oracle RAC 11*g* Release 2 ..... **12**
Current certifications 12
Oracle High Available IP (HA)IP 13
Oracle Clusterware 14
IBM PowerHA (formerly High Availability Cluster Multi-Processing) 15
General Parallel File System 16
Virtualization for Oracle RAC 11*g* Release 2 17
Network interconnects 19
Tuning tips **21**
Automatic Storage Management 21
Asynchronous I/O 22
Concurrent I/O and direct I/O 22
File system cache size 23
RAC IPC 23
General Parallel File System 23
Oracle process memory footprint 23
Tuning resources 23
Summary **25**
Appendix 1: Resources **26**
About the author **27**
Trademarks and special notices **28**

Change history

Version	Date	Editor	Editing description
1.0	2010.08.25	Ravi Shanmugam	Original Version
2.0	2011.03.31	Ravi Shanmugam	Updated for latest certifications
2.1	2011.10.11	Ravi Shanmugam	Updated for HAIP and PSU

Abstract

This paper consolidates the information necessary for planning and implementing Oracle Database 11g Release 2 (11.2) single node database or Oracle Real Application Clusters (RAC) 11g Release 2 on the IBM AIX® operating system.

This paper summarizes the information available at the time of publication. It will be updated as changes (eg. new certifications) occur.

This paper is written to a level of detail that assumes readers have an in-depth knowledge of AIX, Oracle Database 11g Release2, Oracle RAC 11g Release 2 and the related products.

Introduction

There are many technical topics to consider when running Oracle Database 11*g* Release2 and Oracle RAC 11*g* Release 2 on AIX such as: status of Oracle certifications, 11*g* Release 2 patch sets, AIX code levels, tuning and related software components - just to name a few. The documentation for these topics is spread across many websites, documents, presentations and forums. This paper consolidates that information for easy reference.

This paper focuses on AIX 5.3, AIX 6.1 and AIX 7.1, since these are the AIX versions certified on Oracle Database 11*g* Release 2 and Oracle RAC 11*g* Release 2.

This is a companion paper to three other papers.

- For 9*i* and 10*g*R1, *Oracle 9i & 10gR1 on IBM AIX5L: Tips and Considerations*, http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100556
- For 10*g*R2, *Oracle Database 10gR2 and Oracle RAC 10gR2 on IBM AIX: Tips and Considerations*, http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101089
- For 11*g*, *Oracle Database 11g and Oracle RAC 11g on IBM AIX: Tips and Considerations*, http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101176

Feedback is important; please send any comments about this paper to the IBM Oracle International Competency Center at ibmoracl@us.ibm.com.

The IBM Oracle International Competency Center (ICC) works closely with the IBM Oracle Joint Solutions Center (JSC) in Montpellier, France and the IBM Oracle Competency Center in Tokyo, Japan.

- IBM Oracle Center, Montpellier, France – ioc@fr.ibm.com
- IBM Oracle Competency Center, Tokyo, Japan – ibmoracc@jp.ibm.com

Terminology

In 2008, the System p™ and System i™ product families were combined to create the IBM Power™ Systems product family. The IBM Power Systems product family includes systems previously referred to as System p, System p5™, eServer™ pSeries®, RS6000®, System i™, System i5™, eServer iSeries® and AS/400®,.

The brand name of “AIX 5L™” is no longer used. The term “AIX 5L” will still appear in some places in this document, mostly in document titles which have yet to be updated.

High Availability Cluster Multi-Processing (HACMP™) has been renamed to the PowerHA™ and PowerHA SystemMirror. This version of the document will use the name HACMP for version up to 5.4.1, PowerHA for version 5.5 and PowerHA SystemMirror for the version 6.1 and up.

IBM AIX

AIX is an open standards-based, UNIX operating system. AIX in combination with IBM's virtualization offerings provides new levels of flexibility and performance to allow you to consolidate workloads on fewer servers which can increase efficiency and conserve energy. AIX delivers high levels of security, integration, flexibility and reliability—essential for meeting the demands of today's information technology environments. AIX operates on IBM Systems based on Power Architecture technology. For more information about AIX see this web page: http://www-03.ibm.com/systems/p/os/aix/.

IBM’s latest version of AIX is 7.1. It contains new features for virtualization, security, availability and manageability. AIX 7.1 is binary compatible with AIX 6.1 and AIX 5.x. AIX 7.1 includes supporting high end POWER7 server with 256 CPUs, which can provide up to 1024 logical processors with use of SMT 4.

For additional information about AIX 7.1, visit this web page: http://www-03.ibm.com/systems/power/software/aix/v71/index.html

Most of the new features of AIX 7.1 are available on earlier POWER™ platforms. However, the best capabilities are delivered on systems based on POWER6™ and POWER7™ processors. POWER7 processors built for Smarter Planet and automatically optimize work load performance and capacity. POWER7 loaded with specific features such as TurboCore™ for maximum per core performance for databases and MaxCore for high rate of parallelization and high capacity throughput. The POWER6 processor has ultra high-frequency dual-core processor technology and two integrated hardware accelerators. POWER7 processors built for Smarter Planet and automatically optimize work load performance and capacity. For information about the POWER7 and POWER6 processors, see this web page:
http://www-03.ibm.com/systems/power/advantages/power.html.

The following web link has additional information of AIX versions and their Technology Levels (TL) supported on Power System servers.

IBM AIX “From Strength to Strength – A summary of upgrade benefits for each release of AIX”

http://public.dhe.ibm.com/common/ssi/ecm/en/poo03022usen/POO03022USEN.PDF

Large IBM Power Systems Power7 server can have up to 256 cores and supports up to 1024 logical processors.
Power Systems with POWER7 processors requires minimum level of IBM SDK as specified in the following link
http://www.ibm.com/developerworks/java/jdk/power7/index.html

IBM also produces benchmark reports that demonstrate the performance results of Power Systems running AIX. These benchmarks are run using Oracle E-Business Suite, yet they provide valuable data

about Oracle Database and Oracle RAC. The benchmark results are published at http://www.oracle.com/apps_benchmark/html/results.html.

IBM will also provide sizings estimations to help predict the system resources necessary to support a given workload. To start the sizing process, visit this web page, http://www.ibm.com/erp/sizing.

Additional documentation resources for AIX can be found at:

- IBM Power Systems and AIX Information Center, http://publib16.boulder.ibm.com/pseries/index.htm
- IBM developerWorks AIX, http://www-128.ibm.com/developerworks/aix
- IBM AIX Wiki, http://www-941.ibm.com/collaboration/wiki/display/WikiPtype/Home

AIX support

The latest fixes and updates for your system's hardware and operating system can be found at the **Fix Central** web page: http://www-912.ibm.com/eserver/support/fixes/.

Be sure to review My Oracle Support note 282036.1, "Minimum Software Versions and Patches Required to Support Oracle Products on IBM pSeries", for the latest, up-to-date issues regarding AIX and Oracle Database 11*g*R2. This My Oracle Support note is the primary mechanism used to broadcast any breaking news, such as PTF's or bug fixes, about AIX and Oracle Database 11*g*R2.

In general, the ICC recommends that you keep your TL's and SP's up to date for your AIX installation. Monitor My Oracle Support note 282036.1 and Fix Central for the latest issues. Always perform thorough testing on an OS update before deploying to production.

Note: AIX 6.1 TL5 SP1 is recommended for POWER7 for the issue -Threads are not being utilized efficiently in SMT4 and Performance impact observed.

Obtain and Install the eFix IZ78005 for AIX 6.1 TL05 SP1 for the issue - PERFORMANCE DEGRADATION IN IO WHEN AIO BUF SIZE IS > 128K or upgrade to AIX6.1 TL05 SP02 and UP

Obtain the following fixes for AIX 7.1 and install

- IZ87216: UNMAPPING ONE LUN ON DS5300 CAUSES ACCESS TO HANG ON OTHER LUNS
- IZ87564: 7.1. POW() SUBROUTINE IN AIX 7100-00-01-1037, BOS.ADT.LIBM

Recommended Code Levels

There are two tools to help determine the recommended code levels among AIX and Power Systems related components.

1. The Fix Level Recommendation Tool (FLRT) can determine the recommended code levels among a mixture of AIX, HMC, Server firmware, VIOS, GPFS™ and PowerHA. The FLRT web page is: http://www14.software.ibm.com/webapp/set2/flrt/home. Note, the FLRT recommendation provides a minimum acceptable level of compatibility.

2. The POWER code matrix indicates the recommended code levels for the HMC and Server firmware. The POWER code matrix web page is: http://www14.software.ibm.com/webapp/set2/sas/f/power5cm/home.html. Note, the POWER code matrix recommendation provides the maximum stability code combinations.

Service Strategy

To review the latest *IBM AIX Operating System Service Strategy Details and Best Practices* document, see this website: http://www14.software.ibm.com/webapp/set2/sas/f/best/home.html.

C and C++ compilers

XL C/C++ is a standards-based, high performance C and C++ compiler with advanced optimizing and debugging features. It gives you the ability to optimize and tune applications for optimal execution on systems using PowerPC®, POWER3™, POWER4™, POWER5™, POWER6™ and the latest POWER7™ processors from IBM. The compiler supports IBM Power Systems servers capable of running IBM AIX 5.2, 5.3, 6.1 and 7.1.

To determine the ***current certifications*** for compilers on 11*g* Release 2, please review My Oracle Support note 43208.1, "Certified Compilers". Currently, XL C/C++ 8.0 and 9.0 are certified for Oracle Database 11*g*. If necessary, read My Oracle Support note 335569.1, "How to Find C or C++ version on AIX Platform" to determine your XL C/C++ compiler version.

The XL C/C++ runtime environment is installed during the installation of base AIX. To update to the latest runtime environment, go to the XL C/C++ web page, http://www-306.ibm.com/software/awdtools/xlcpp/features/aix/ and click on the "Latest XL C/C++ Updates (PTFs)" link. Then select the "Runtime Environment Components" in the AIX section.

If the XL C/C++ Enterprise Edition for AIX compiler is installed, to get the latest updates, visit the product's web page, http://www-306.ibm.com/software/awdtools/xlcpp/features/aix/ and click on the "Latest XL C/C++ Updates (PTFs)" link and select the appropriate link in the AIX section.

Oracle Database 11*g* Release 2

This section contains the Oracle Database 11*g* non-RAC technical information that needs to be considered in an AIX installation. Oracle 11g Release 2 introduces the Grid Infrastructure installation for both stand alone and Real Application Cluster (RAC) databases. Grid Infrastructure is an option which includes Automatic Storage Management (ASM), the listener, and oracle Restart. Oracle Restart is a new feature that provides the ability to monitor, manage, and automatically restart the Oracle database instance, ASM instance, and listeners.

Grid infrastructure is needed for the stand alone database server if the ASM is selected as a storage option for the database files.

These are the basic documents to review for an Oracle Database 11*g* installation on AIX.

- "Oracle Grid Infrastructure Installation Guide 11*g* Release 2 (11.2) for IBM AIX Based Systems E10814-03 http://download.oracle.com/docs/cd/E11882_01/install.112/e10814/toc.htm
- "Oracle Database Installation Guide 11*g* Release 2 for AIX on POWER Systems ", E10854-05, http://download.oracle.com/docs/cd/E11882_01/install.112/e10854/toc.htm

- "Oracle Database Release Notes 11*g* Release 2 (11.2) for AIX on POWER Systems", E10853-03, http://download.oracle.com/docs/cd/E11882_01/relnotes.112/e10853/toc.htm
- My Oracle Support note 282036.1, "Minimum Software Versions and Patches Required to Support Oracle Products on IBM pSeries" for the latest, up-to-date issues regarding Oracle Database 11*g* Release 2 and AIX.
- Oracle Database Editions and Options, http://www.oracle.com/database/product_editions.html

Note: Even though the first two Oracle documents referenced above indicate "AIX 5L" in the title, the content applies equally well to AIX 6.1 and AIX 7.1

Current certifications

To determine the ***current certifications*** for Oracle Database 11*g* Release 2 on AIX,

- Sign into Oracle's My Oracle Support website (UserID/Password is required) at https://support.oracle.com/CSP/ui/flash.html and click on "Certifications" tab, look for "Certification Search" section. In the "Product" field. type "oracle Database", the second field is "Release". Select one of them from the list. It has the releases from 8.1.7.4 to 11.2.0.2.0. next field is "Platform", select "IBM AIX on POWER Systems", it lists the AIX versions from 5.1 to 6.1.Choose one of them. Then click "Search" button. The search result will show the certification status in the left side as well as any AIX critical patch and support information in the right side.

This lists the information with Certification status for the above input.

The following table shows the high level certification information.

OS	Product	Status
AIX 5.3 TL09 SP01 or later AIX 6.1 TL02 SP01 or later AIX 7.1 TL0 SP01 or later	11*g* Release 2 64-bit	Certified

Table 1: Oracle Database Enterprise Edition 11g Release 2 certifications of AIX as of publication date.

The same certifications are in place for Standard Edition (SE) as Enterprise Edition (EE).

Here are some certification details to be aware of:

- These products are certified for AIX 5.3, 6.1 and 7.1 on all Power Systems servers supported by those versions of AIX.
- 64-bit hardware is required for a 64-bit application such as Oracle Database 11*g* for AIX.
- Servers capable of more than 4 processors are certified only for EE.
- Logical Partitioning (LPARs), Dynamic Logical Partitioning (DLPARs) and Micro-Partitioning™ are supported.
- Virtual IO Server (VIOS) is supported.
- Apply AIX 7.1 fix for APAR IZ87564 and IZ87216

Latest patch set

At the time of publication the base and the latest patch set is 11.2.0.2. To find the documentation for the latest Oracle Database 11*g* Release 2 patch set, sign into Oracle's My Oracle Support website at https://support.oracle.com/CSP/ui/flash.html. In the Quick Find pulldown, specify "Document ID". Then specify Document Number 161818.1, "Oracle Server (RDBMS) Releases Support Status Summary". Then click on "11.2.0.X", this will bring up My Oracle Support note 880782.1, "Support Status and Alerts for Oracle 11*g* Release 2(11.2.0.X)". Then go to the "IBM AIX Based Systems (64-bit)" section and locate the ReadMe (if any) for 11.2.0.2. The patch number is 10098816 and the files (7 files) are p10098816_AIX64_5L_1of7.zip …. p10098816_112020_AIX64-5L_7of7.zip.

Note, even though the zip file to download may contain "AIX5L" in the name, the download is for AIX 5.3, 6.1 and 7.1.

Note: Important Changes to Oracle Database Patch Sets Starting With 11.2.0.2, My Oracle Support note 1189783.1

> "Starting with the first patch set for Oracle Database 11g Release 2 (11.2.0.2), Oracle Database patch sets are full installations of the Oracle Database software. In past releases, Oracle Database patch sets consisted of a set of files that replaced files in an existing Oracle home. Beginning with Oracle Database 11g Release 2, patch sets are full installations that replace existing installations."

Please read the full My Oracle Support note for the detailed information.

Also, check the Oracle Database 11*g* Release 2 release notes for AIX Based Systems, http://download.oracle.com/docs/cd/B28359_01/relnotes.111/b32075/toc.htm, for additional AIX-specific information.

In general, the ICC recommends that you keep your patch set current for your Oracle Database 11*g* Release 2 installation. Always perform thorough testing on a database update before deploying to production. Monitor the My Oracle Support note 282036.1 and the Release Notes for AIX Based Systems for the latest issues.

Critical Patch Update schedule

The Critical Patch Update (CPU) program is the method by which Oracle delivers security patch updates and security fixes for all their products. A CPU is a collection of patches for multiple security vulnerabilities. It also includes non-security fixes that are required (because of interdependencies) by those security patches. Oracle provides CPUs for all product offerings on a quarterly schedule.

For more information about CPUs, read the page at http://www.oracle.com/technology/deploy/security/cpu/cpufaq.htm - "Security Alerts and Critical Patch Updates – Frequently Asked Questions". Watch for Critical Patch Updates at this site: http://www.oracle.com/technology/deploy/security/alerts.htm.

Patch Set Update (PSU)

Beginning with October 2009 Critical Patch Update release, Oracle delivers Patch Set Updates for all platforms on the release date. PSUs are proactive cumulative patches containing recommended bug fixes that are released on a regular and predictable schedule. PSUs are on the same quarterly

schedules as the Critical Patch Update (CPU). PSUs consist of CPU, generic patch bundles, RAC patch bundle, and Data Guard patch bundles.

The current Patch Set Update at the time of publication of this document was 11.2.0.2.3 for database Patch number for the update of Database and Grid Infrastructure (GI) was 12419353. See My Oracle Support note 1323616.1 “Patch Set Update and Critical patch Update July 2011 Availability Document” and look for “oracle Database Patch Set Update” and download the Patch 12419353 “p12419353_112020_AIX64-5L.zip”.

If the customers did not update their database version to 11.2.0.2 and staying with 11.2.0.1.x, the Database PSU is 11.2.0.1.6. The patch number is 12419378 “p12419378_112010_AIX64-5L.zip”.

Virtualization for Oracle Database 11*g* Release 2

As mentioned in the Current Certifications sub-section (see above) for Oracle Database 11*g* Release 2; LPARs, Micro-Partitioning, Dynamic LPAR and VIOS are supported. LPARs, Micro-Partitioning, Dynamic LPAR and VIOS are part of PowerVM™. PowerVM is the family of technologies, capabilities and offerings that deliver industry-leading virtualization on IBM POWER processor-based systems. The PowerVM capabilities supported in Oracle Database 11*g* Release 2 are:

- **Logical Partitions** subdivide a computer's processors, memory, and hardware resources into multiple environments so that each environment can be operated independently with its own operating system and applications.
  - **Dedicated processor partitions** are LPARs that use dedicated processors.
    - **Dedicated processors** are whole physical processors that are assigned to a single LPAR.
  - **Shared Processor partitions** are LPARs that use Micro-Partitioning in conjunction with a shared processor pool.
    - **Micro-Partitioning** divides a physical processor’s computing power into fractions of a processing unit and shares them among logical partitions. Processing capacity can be configured in fractions of 1/100 of a processor. The minimum amount of processing capacity that has to be assigned to a partition is 1/10 of a processor.
    - A **shared processor pool** is a group of physical processors that are not dedicated to any LPAR.
- **Dynamic LPARs** (DLPAR) are a shared or dedicated LPAR to which changes can be made to the amount of processors, memory and virtual or physical adapters without requiring a reboot.
  - For dedicated processor partitions, it is only possible to dynamically add, move or remove whole processors. When a processor is removed from a dedicated processor partition, it is then assigned to the shared processor pool.
  - For a shared processor partition, it is also possible to dynamically change the shared processor capacity, the weight of the uncapped attribute, virtual processors and capped/uncapped mode.
- **Virtual I/O Server** allows sharing of physical resources between logical partitions (LPARs) including virtual SCSI and virtual networking. This allows more efficient utilization of physical resources through sharing between LPARs and facilitates server consolidation.

  VIOS 2.1 (2.1.2.10. FP 22) is certified on Oracle 11*g* R2 (11.2.0.1) with AIX 6.1 TL03 SP01.

VIOS 2.2 (2.20.10.FP24) with APAR IZ86878 is supported on AIX 7.1 11*g*R2 with AIX 7.1 along with fixes for APAR IZ87564 and IZ87216 or later.

With certified VIOS combinations customers may use virtual SCSI with or without "N Port ID virtualization (NPIV) to attach disk for data storage and associated raw disk-based voting or OCR.This may be done for both ASM and GPFS.
The VIOS Support web page is http://www14.software.ibm.com/webapp/set2/sas/f/vios/home.html

- **Workload Partition** (WPAR) is a software-based virtualization feature. WPARs subdivide an AIX instance into multiple environments, each hosting applications and providing isolation from applications executing in other environments. **Live Application Mobility** allows you to relocate running WPARs from one LPAR to another. For more information on WPARs, see the *Introduction to Workload Partition Management in IBM AIX Version 6.1* Redbook, http://www.redbooks.ibm.com/abstracts/sg247431.html?Open

  Oracle DB 11*g*R2 (11.2.0.2) is certified to use with WPAR in AIX 6.1 TL02 SP02..
  Apply AIX patch IZ52319 and IZ54871 and set "export AIXPERFSTAT_SCOPE=M"
  - Only IBM AIX JFS2 and supported NAS devices can be used with WPAR. Oracle's ASM is not supported at this time. Live Application Mobility of WPAR is not certified with Oracle at this time.

PowerVM feature Live Partition Mobility **still being evaluated** for support on Oracle 11*g* Release 2

- **Live Partition Mobility** allows you to migrate running AIX and Linux LPARs and their hosted applications from one physical server to another without disrupting infrastructure services. The migration transfers the entire partition state, including the processor context, memory, attached virtual devices, and connected users. For more information on LPM, see the *PowerVM Live Partition Mobility on IBM System p* Redbook, http://www.redbooks.ibm.com/abstracts/sg247460.html?Open

Oracle DB 11*g* R2 (11.2.0.1) (Single Instance) is certified to use with IBM PowerVM LPM on the following code levels or higher are recommended

- AIX 6.1 TL4 SP01
- Oracle Database 11.2.0.1
- VIOS 2.1.2.10.FP22
- HMC V7R7.2.0
- Firmware E*340_122

Customers should monitor the latest Oracle Certification information to be aware of the ongoing certification of new features.

For additional, detailed PowerVM information, see these documents and websites:
http://www.oracle.com/technetwork/database/virtualizationmatrix-172995.html

- PowerVM webpage, http://www-03.ibm.com/systems/power/software/virtualization/index.html
- PowerVM Virtualization on IBM System p: Introduction and Configuration Fourth Edition, http://www.redbooks.ibm.com/abstracts/sg247940.html?Open

- IBM System p PowerVM Best Practices, http://www.redbooks.ibm.com/abstracts/redp4194.html?Open

Oracle RAC 11*g* Release 2

This section contains the Oracle RAC 11*g* technical information that needs to be considered in an AIX installation. In Oracle 11*g* RAC Release 2, Grid Infrastructure includes Oracle Clusterware, Automatic Storage Management and the listener. These are the basic documents to review for an Oracle RAC 11*g* Release 2 installation on AIX.

- "Oracle Grid Infrastructure Installation Guide 11*g* Release 2 (11.2) for IBM AIX Based Systems E10814-03 http://download.oracle.com/docs/cd/E11882_01/install.112/e10814/toc.htm
- "Oracle Real Application Clusters Installation Guide 11*g* Release 2 (11.2) for Linux and UNIX",E10813-06. http://download.oracle.com/docs/cd/E11882_01/install.112/e10813/toc.htm
- My Oracle Support note 282036.1, "Minimum Software Versions and Patches Required to Support Oracle Products on IBM pSeries" for the latest, up-to-date issues regarding Oracle Database 11*g* and AIX.

Note, these documents apply equally well to AIX 5L version 5.3, AIX 6.1 and 7.1.

Current certifications

To determine the ***current certifications*** for Oracle RAC 11*g* Release 2 on AIX,

- Sign into Oracle's My Oracle Support website (UserID and Password required) at https://support.oracle.com/CSP/ui/flash.html and click on "More.." and click "Certifications" tab, look for "Certification Search" section. In the "Product" field. type "oracle Database", the second field is "Release". Select one of them from the list. It has the releases from 8.1.7.4 to 11.2.0.2.0. next field is "Platform", select "IBM AIX on POWER Systems", it lists the AIX versions from 5.1 to 7.1.Choose one of them. Then click "Search" button. The search result will show the certification status in the left side as well as any AIX critical patch and support information in the right side.

. This lists the information with Certification status for the above input.

The following table shows high level information on the certification of Oracle 11*g*R2 with AIX.

OS	Product	Certified With	Version	Status
AIX 5.3,6.1 and 7.1	11*g*R2 64-bit	Oracle Clusterware	11*g*	Certified
AIX 5.3 and 6.1	11*g*R2 64-bit	PowerHA SystemMirror * *(This is for raw logical volume support)*	6.1	Certified
AIX 5.3 and 6.1	11*g*R2 64-bit	HACMP *This is for raw logical volume support. See PowerHA sub-section below.*	5.4.1 SP1	Certified

AIX 5.3 and 6.1	11*g*R2 64-bit	*IBM General Parallel File System (GPFS)*	3.3	Certified
AIX 5.3 and 6.1	11*g*R2 64-bit	*IBM General Parallel File System (GPFS)*	3.2	Certified
AIX 5.3 and 6.1	11gR2 64bit	*Veritas Storage Foundation for Oracle RAC*	5.0	Certified

Table 2: Oracle RAC 11g Release 2 *certifications of AIX as of publication date.*

* IBM HACMP is renamed to PowerHA up to version 5.5, from 6.1, it is called PowerHA SystemMirror.

Here are some certification details to be aware of:

- These products are certified for AIX 5.3,AIX 6.1 and 7.1 on all Power Systems servers supported by those versions of AIX.
- AIX on System i partitions are also supported.
- 64-bit hardware is required for Oracle RAC 11*g* Release 2.
- AIX 64-bit is required for Oracle RAC 11*g* Release 2.
- Logical Partitioning (LPARs), Dynamic Logical Partitioning (DLPARs) and Micro-Partitioning are supported in Oracle RAC environments.
- Virtual IO Server (VIOS) features:
  - See the Virtualization sub-section below for details.

By following the navigation instructions in the first paragraph of this sub-section, the RAC Technologies Compatibility Matrix will also be visible. This matrix supplies details about storage technologies, network interconnect technologies and other platform-specific information.

http://www.oracle.com/technology/products/database/clustering/certify/tech_generic_unix_new.html

Oracle High Available IP (HA)IP

With Oracle Database 11g Release 2, Patch Set One (11.2.0.2), Oracle introduced an integrated Redundant Interconnect Usage feature, which provides a highly available (HA)IP network functionality for the Oracle interconnect. Previous to this version, Oracle RAC and Oracle Clusterware depended on AIX and respective OS features to provide a highly available network interface for the Oracle interconnect. With Oracle Database 11g Release 2 Patch Set One, customers have the choice to either continue to use the AIX provided HA network interface, or to use Oracle's integrated Redundant Interconnect Usage feature, which will provide full high availability for an Oracle RAC Database and Oracle ASM of version 11.2.0.2 or higher. Oracle's Redundant Interconnect Usage feature will protect production RAC databases where instances of the same database are not co-located on the same physical frame.

For upgrade customers, it is recommended to maintain their current, typically Etherchannel based, configuration as with pre-11.2.0.2 releases during upgrades. This will allow the Redundant Interconnect Usage to allocate an (HA)IP on top of the Etherchannel device, but will not enable load balancing or network failover based on the Oracle Redundant Interconnect Usage feature. Load balancing as well as network failover will continue to be managed by Etherchannel in this case; no further configuration steps required.

In order to fully enable Redundant Interconnect Usage to manage load balancing and network failover for

the Oracle cluster interconnect, the Etherchannel configuration used for the Oracle interconnect should be removed and Oracle Redundant Interconnect Usage should be enabled directly on the devices formerly managed by Etherchannel. For more information refer to the Oracle Documentation on how to enable Redundant Interconnect Usage.

Hosting more than one instance of a production Oracle RAC in the same physical environment or frame with a single point of failure (sharing components required for network connectivity, storage access, common Hypervisor, or other critical components) at the same time is generally not recommended by Oracle for a complete High Availability solution, as a failure of any of those shared components inevitably affects more than one instance of the production Oracle RAC database. Under certain circumstances, virtualization solutions and other techniques provided by the hardware or OS vendor may mitigate these negative effects, however, for critical and production deployments, clustering within the same frame, if it has a single point of failure, is discouraged. Furthermore, at this point in time, when enabling Redundant Interconnect Usage, avoid co-location of Oracle RAC instances belonging to the same production database on the same frame as described above, when configured with virtual Ethernet, as certain failures (e.g. the loss of a physical network and one VIO server) in the frame could lead to losing the majority of the Oracle RAC database instances. Oracle and IBM are working to integrate the Redundant Interconnect Usage feature so that optimized high availability can be ensured. Alternatively, physical devices (as opposed to virtual or VIO based devices) can be used and managed by the Redundant Interconnect Usage feature directly to avoid such scenarios.

Oracle Clusterware

In a RAC environment Oracle Clusterware provides the high availability functionality. This includes monitoring or restarting the nodes of the cluster, for the database instances, for the listeners and for the database services. Oracle Clusterware is required for Oracle RAC 11*g* Release 2. For more information on Oracle Clusterware, visit: http://www.oracle.com/technology/products/database/clustering/index.html.

For AIX 6.1, to resolve Oracle Bug 7006789 “VIP does not failover when public network cables are removed on one node”, follow this configuration recommendation.

> Oracle RAC clusterware has strict timeout requirements for VIP address failover in case of a public network failure. When DNS servers are unreachable due to a public network failure, DNS name resolution calls such as getaddrinfo may hang for the default AIX query timeout duration of 5 minutes. Name resolution calls made by Oracle processes can thus delay the VIP failover. To reduce such delays, the DNS query timeout can be reduced to 1 minute, by adding the following options line in /etc/resolv.conf for all RAC cluster nodes:
>
> "options timeout:1"
>
> No reboot is necessary to activate this change.
>
> If you need even faster VIP failover the timeout can be further reduced to a value of 0; provided your network infrastructure (network and DNS servers) has the speed to serve name queries within a few (5-6) seconds. If you use a value of 0 for timeout and your DNS or network is slow to respond, DNS name lookups will start to fail prematurely.

IBM PowerHA (formerly High Availability Cluster Multi-Processing)

Note: High Availability Cluster Multi-Processing (HACMP) has been renamed to the PowerHA Cluster Manager. This version of the document will use the term PowerHA.

For Oracle RAC 11*g* Release 2, raw logical volumes are only supported when performing an upgrade installation from oracle RAC 10*g* R2 or 11*g*R1. Raw logical volumes are not available when performing a fresh install of Oracle RAC 11*g* Release 2.

Where customers desire raw logical volumes for their Clusterware files, database files or recovery files, PowerHA will be required. The certification details for the different versions of PowerHA on Oracle RAC 11*g*R2 are:

HACMP 5.4.1

- **Certified on AIX 5.3 and AIX 6.1.**

The certification notes for HACMP 5.4.1 and **AIX 5.3** list the following requirements.

- AIX 5.3 TL 6 or later, specifically bos.rte.lvm 5.3.0.60. 11*g* Release 2 requires TL09 SP01 or later.
- HACMP 5.4.1, available in media or APAR IZ02620.
- RSCT 2.4.5 for AIX 5.3
- Reliable Scalable Cluster Technology (RSCT) 2.4.7.3, rsct.basic.rte and ifix for APAR IZ01838. This APAR is incorporated into 2.4.8.1.
- Patch 6718715 for rootpre.sh is required with HACMP during a fresh install of Oracle RAC Clusterware or when upgrading to Oracle Database 11*g* Release 2.

The certification notes for HACMP 5.4.1 and **AIX 6.1** list the following requirements.

- AIX 6.1 TL02 SP1 or later.
- HACMP 5.4.1 SP01, available in media or APAR IZ02620.
- RSCT 2.5.2 for AIX 6.1
- Patch 6718715 for rootpre.sh is required with PowerHA during a fresh install of Oracle RAC Clusterware or when upgrading to Oracle Database 11*g* Release 2.
- To resolve Oracle Bug 7006789 “VIP does not failover when public when public network cables are removed on one node”, follow this configuration recommendation.

  Oracle RAC clusterware has strict timeout requirements for VIP address failover in case of a public network failure. When DNS servers are unreachable due to a public network failure, DNS name resolution calls such as getaddrinfo may hang for the default AIX query timeout duration of 5 minutes. Name resolution calls made by Oracle processes can thus delay the VIP failover. To reduce such delays, the DNS query timeout can be reduced to 1 minute, by adding the following options line in /etc/resolv.conf for all RAC cluster nodes:

  "options timeout:1"

  No reboot is necessary to activate this change.

If you need even faster VIP failover the timeout can be further reduced to a value of 0; provided your network infrastructure (network and DNS servers) has the speed to serve name queries within a few (5-6) seconds. If you use a value of 0 for timeout and your DNS or network is slow to respond, DNS name lookups will start to fail prematurely.

PowerHA SystemMirror 6.1

Oracle RAC 11.2.0.1 is certified with PowerHA SystemMirror 6.1 with following software stack

- AIX 5.3 TL10 SP05 or later, must include fix for APAR IZ89268
- AIX 6.1 TL04 SP07 or later, must include fix for APAR IZ87768
- VIOS 2.2.0.10 FP24 SP01
- PowerHA 6.1 SP03 or later
- RSCT 2.4.12.0 or later for AIX 5.3
- RSCT 2.5.4.0 or later for AIX 6.1

For more information, refer the My Oracle Support note 404474.1 “Status of Certification of Oracle Clusterware with PowerHA” and read the content in the topic “Using PowerHA and Oracle RAC 11*g*R2”.

PowerHA version compatibility matrix with AIX versions can be found in the link

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD101347

Multi-node Disk Heartbeat (MNDHB)

In order to reduce the likelihood of an unnecessary cluster partition failover, multiple IP networks and at least one non-IP network are recommended for heart beating. The most convenient way of configuring non-IP networks is to use disk heart beating, as it removes the problems of distance associated with RS-232 networks. Refer to these documents for additional information on MNDHB:

- PowerHA Best Practices, http://www-03.ibm.com/systems/clusters/whitepapers/hacmp_bestpractices.html
- My Oracle Support note 404474.1, “Status and Certification of Oracle Clusterware with HACMP 5.3 & 5.4” for detailed MNDHB configuration instructions.

The PowerHA web page is at http://www-03.ibm.com/systems/p/ha/.

General Parallel File System

If using a file system for your Oracle RAC 11*g* Release 2 data files (as opposed to raw logical volumes or ASM), you’ll need to use a cluster file system (CFS). A CFS allows file system access by all members in a cluster at the same time. That requirement precludes JFS and JFS2 from being used for Oracle RAC 11*g* data files. The IBM General Purpose File System (GPFS) is an Oracle RAC 11*g* certified CFS.

GPFS is a high-performance shared-disk file system that can provide fast, reliable data access from all nodes in a homogenous or heterogeneous cluster of IBM UNIX servers running either the AIX or the Linux® operating system.

To determine the ***current certifications*** for GPFS on Oracle RAC, browse to the "General Notes for RAC for Unix on IBM AIX based Systems (RAC only)" web page, as described in the beginning of the Current Certifications sub-section (see above).

GPFS 3.3 is certified with Oracle 11*g*R2 with AIX 5.3 and 6.1. The following stacks are used for certification

- GPFS 3.3 (fixpack 3.3.0.6)
- Oracle RAC 11.2.0.1
- AIX 6.1 TL04 SP04
- VIOS 2.1 with IBM NPIV (N-Port ID Virtualization)

GPFS 3.2 is certified with Oracle 11*g*R2 with AIX 6.1. The following are the certified code level. Use the following code level or later.

- GPFS 3.2.1.14
- Oracle RAC 11.2.0.1
- AIX 6.1 TL03 SP01
- VIOS 2.1.0.10.FP 20.1

Refer the My Oracle support note 302806.1 for more information on GPFS for Oracle RAC.

The GPFS web site is http://www-03.ibm.com/systems/clusters/software/gpfs.html. Make sure to review the ***current GPFS advisories*** in the GPFS FAQ available from the GPFS web site.

See the GPFS sub-section in the Tuning Tips section (below) for GPFS tuning information.

Virtualization for Oracle RAC 11*g* Release 2

As mentioned in the Current Certifications sub-section (see above) for Oracle RAC 11*g* Release 2; LPARs, Micro-Partitioning and Dynamic LPAR are supported in Oracle RAC 11*g* Release 2 environments. LPARs, Micro-Partitioning, Dynamic LPAR and VIOS are part of PowerVM. PowerVM is the family of technologies, capabilities and offerings that deliver industry-leading virtualization on IBM POWER processor-based systems. The PowerVM capabilities supported in Oracle RAC 11*g* Release 2 are:

- **Logical Partitions** subdivide a computer's processors, memory, and hardware resources into multiple environments so that each environment can be operated independently with its own operating system and applications.
  - **Dedicated processor partitions** are LPARs that use dedicated processors.
    - **Dedicated processors** are whole processors that are assigned to a single LPAR.
  - **Shared Processor partitions** are LPARs that use Micro-Partitioning in conjunction with a shared processor pool.
    - **Micro-Partitioning** divides a physical processor's computing power into fractions of a processing unit and shares them among logical partitions. Processing capacity can be configured in fractions of 1/100 of a processor. The minimum amount of processing capacity that has to be assigned to a partition is 1/10 of a processor.

- A **shared processor pool** is a group of physical processors that are not dedicated to any LPAR.

- **Dynamic LPARs** (DLPAR) are a shared or dedicated LPAR to which changes can be made to the amount of processors, memory and virtual or physical adapters without requiring a reboot.
  - For dedicated processor partitions, it is only possible to dynamically add, move or remove whole processors. When a processor is removed from a dedicated processor partition, it is then assigned to the shared processor pool.
  - For a shared processor partition, it is also possible to dynamically change the shared processor capacity, the weight of the uncapped attribute, virtual processors and capped/uncapped mode.
- **Virtual I/O Server** allows sharing of physical resources between logical partitions (LPARs) including virtual SCSI and virtual networking. This allows more efficient utilization of physical resources through sharing between LPARs and facilitates server consolidation. These VIOS features are specifically supported in Oracle RAC 11*g* Release 2:
  - Virtual LAN for public and private interconnects and all supported data storage options.
  - Two VIO servers are necessary to improve serviceability and provide the availability required for an Oracle RAC 11*g* Release 2 implementation.
  - VIOS can be used for non-RAC functions such as Virtual SCSI based root volume groups (rootvg), and Virtual SCSI & NPIV (N port ID Virtualization) for ASM data storage.
  - VIOS with ASM and GPFS in Oracle RAC 11*g* Release 2 is supported

  In VIOS environment, NPIV feature allows multiple N_Port IDs to share a single Physical N_Port, which means multiple client LPARs.can access external storage LUNs through the same Fibre channel adapter assigned to VIOS LPAR.

  Note: With certified VIOS combinations customers may use either Virtual SCSI (vSCSI) or N Port Id Virtualization (NPIV) to attach disk for data storage and associated raw disk based Voting and OCR. This may be done for both ASM and GPFS 3.2. Customer must confirm IBM support of the configuration, and install any required AIX and Oracle updates before using.

VIOS 2.1 and VIOS 2.2 are certified on Oracle RAC 11*g*R2 with AIX 6.1 and 7.1 respectively

VIOS2.2:

- VIOS 2.20.10.FP24 with APAR IZ86878
- Oracle RAC 11.2.0.1
- AIX 7.1 SP01 with fixes for APAR IZ87564 and IZ87216 or later

VIOS 2.1:

- VIOS 2.1.2.10.FP22
- Oracle RAC 11.2.0.1
- AIX 6.1 TL03 SP1 or later

The VIOS Support web page is:
http://www14.software.ibm.com/webapp/set2/sas/f/vios/home.html.

These are some PowerVM features **still being evaluated** for support on Oracle RAC 11*g* Release 2:

- **Workload Partition** (WPAR) is a software-based virtualization feature. WPARs subdivide an AIX instance into multiple environments, each hosting applications and providing isolation from applications executing in other environments. **Live Application Mobility** allows you to relocate running WPARs from one LPAR to another. For more information on WPARs, see the *Introduction to Workload Partition Management in IBM AIX Version 6.1* Redbook, http://www.redbooks.ibm.com/abstracts/sg247431.html?Open
- **Live Partition Mobility** allows you to migrate running AIX and Linux LPARs and their hosted applications from one physical server to another without disrupting infrastructure services. The migration transfers the entire partition state, including the processor context, memory, attached virtual devices, and connected users.

For more information on LPM, see the *PowerVM Live Partition Mobility on IBM System p* Redbook, http://www.redbooks.ibm.com/abstracts/sg247460.html?Open

Customers should monitor the latest Oracle Certification information to be aware of the ongoing certification of new features.

Virtualization features are powerful, yet their implementation can get quite complex in Oracle RAC 11*g* Release 2. If you have any questions about implementing virtualization with Oracle RAC 11*g* Release 2, contact the IBM Oracle International Competency Center at ibmoracl@us.ibm.com. For additional, detailed PowerVM information, see these documents and websites:

- PowerVM webpage, http://www-03.ibm.com/systems/power/software/virtualization/index.html
- PowerVM Virtualization on IBM System p: Introduction and Configuration Fourth Edition, http://www.redbooks.ibm.com/abstracts/sg247940.html?Open
- IBM System p PowerVM Best Practices, http://www.redbooks.ibm.com/abstracts/redp4194.html?Open

Network interconnects

To determine the ***current certifications*** for network interconnects on Oracle RAC, browse to the "General Notes for RAC for Unix on IBM AIX based Systems (RAC only)" web page, as described in the beginning of the Current Certifications sub-section (see above) for Oracle RAC 11*g* Release 2. Then select the link for the RAC Technologies Compatibility Matrix (RTCM). Within RTCM, Network Interconnect certifications are listed. Currently, the following are supported and certified with AIX 5.3, AIX 6.1 and 7.1.

- 100 Mbps ,1 Gigabit
- 10 Gigabit Ethernet
- IP over InfiniBand (IPoIB)

For information about implementing IPoIB in AIX, refer the following documents:

- Implementing InfiniBand on IBM System p (Redbook), www.redbooks.ibm.com/abstracts/sg247351.html?Open
- Configuring InfiniBand for AIX (article), http://www.ibm.com/developerworks/aix/library/au-infiniband/?S_TACT=105AGY06&

The new HAIP feature introduced in 11.2.0.2 is not compatible with InfiniBand on AIX at this time. Customers can continue to use 11gR1 with IB on AIX

Integrated Virtual Ethernet

The Integrated Virtual Ethernet (IVE) is a collection of POWER6 hardware, software and hypervisor features that provides integrated high-speed Ethernet adapter ports with virtualization capabilities. The IVE appears in AIX system commands as the Host Ethernet Adapter (HEA) or Local HEA (LHEA). The IVE offers:

- IVE Adapter Ethernet port options:
  - Two 1 Gbps ports or
  - Four 1 Gbps ports or
  - Two 10 Gbps ports
- External network connectivity for LPARs using dedicated ports without the need of a VIOS.
- Industry standard hardware acceleration, loaded with flexible configuration possibilities.
- The speed and performance of the GX+ bus, faster than PCI Express x16.

For additional information about IVE/HEA:

- Integrated Virtual Ethernet Adapter, Technical Overview and Introduction (Redbook), http://www.redbooks.ibm.com/redpapers/pdfs/redp4340.pdf

Ensure your IVE/HEA is deployed consistent with the instructions in My Oracle Support note 282036.1, “Minimum Software Versions and Patches Required to Support Oracle Products on IBM pSeries”.

Refer the IBM technical document for “Setting up IBM POWER6 10 Gigabit Ethernet ports and AIX 6.1 etherchannel for Oracle RAC private interconnectivity” , Document ID : WP101734

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101734

Some additional consideration as part of the 10 GigE setup as follows,

- LACP timeout: Use the “long timeout” switch setting for the amount of time to wait before sending LACPDUs.
- Flow control: Enable flow control at the switch port and on the server side ports (using HMC) for the 10GE adapter or 10GE HEA configuration.
- UDP tuning: The RAC interconnect uses UDP for interprocess communications. Tune the udp_sendspace and udp_recvspace parameters until the “netstat –s” command indicates there are no “socket buffer overflows”. See the Network Tuning section of the “Oracle Architecture and Tuning on AIX” document:

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100883
- Jumbo frames: Enable Jumbo frames on the RAC interconnect: Configure Jumbo frames at the switch port. In the certification project we set Jumbo frames to 9252 at the switch.

Configure Jumbo frames on the server side ports (using HMC) for the 10GE adapter or 10GE HEA configuration. Configure Jumbo frames in the Etherchannel (network interface) definition.

Note: When using Gigabit Ethernet, 10 Gigabit Ethernet and IP over InfiniBand customers may configure the network routing using either EtherChannel or AIX VIPA based on their requirements. Customer must confirm IBM support of the configuration, and install any required AIX and Oracle updates before using.

Important Oracle fixes or issues not specific to AIX release

Bug 6782569: “11g install AIX VIPS won’t start”. This fix is required when using the IBM Logical Host Ethernet Adapter (LHEA) for the Oracle Public or VIP interfaces.

Bug 6319685: LGWR posts do not scale on some platforms. LGWR posts do not scale well giving longer than needed latency on “log file sync” waits (which involve posting LGWR).

In order to use AIX VIPA, users need to apply patch 10203061, this is also contains in 11.2.0.2.1 patch set.

Apply the Oracle fix for the bug 9457492: OCR DISKGROUP WAS DISMOUNT WHILE I/O ERROR OCCURED ON ONE NODE

Contact Oracle support for the interim/one off fix.

Tuning tips

Tuning Oracle Database 11*g* Release 2 and Oracle RAC 11*g* Release 2 for AIX is documented in detail in many documents. The tuning items chosen for inclusion in this section are those where the defaults are sub-optimal, or settings that require special consideration. Instead of repeating the existing documentation, this section will consolidate references to the existing documentation. The URLs of the documents referenced in this section are listed at the end of the section.

Automatic Storage Management

Automatic Storage Management (ASM) is a feature in Oracle Database 11*g* Release 2 that provides the database administrator with a simple storage management interface that is consistent across all server and storage platforms.

Starting with Oracle Database 11*g* Release 2, Oracle Clusterware OCR and voting disk files can be stored in Oracle ASM disk group.

In Oracle Database 11*g* Release 2 , ASM becomes a complete storage management solution for both Oracle Database and non-database files and has many extended functions for not only storing database files, but also storing binary files, report files, trace files, alert logs and other application data files.

ASM Cluster File Systems (ACFS) extends ASM by providing cluster file system scaled to a large number of nodes and uses extend-based storage allocation for improved performance. ACFS can be can be exported to remote clients through NFS and CIFS.

ASM Dynamic Volume Manager (DVM), ASM FS Snapshot, ASM Intelligent Data Placement, ASM Storage Management Configuration Assistant (ASMCA), ASM File Access Control and ASMCMD are some of the extended functions of ASM.

For more information on ASM new features , refer to the Oracle document "*Oracle Database New Features Guide 11g Release 2 (11.2)"*
http://download.oracle.com/docs/cd/E11882_01/server.112/e10881/chapter1.htm#insertedID0

For information about ASM, see Oracle's ASM web page:
http://www.oracle.com/technology/products/database/asm/index.html

Asynchronous I/O

Asynchronous I/O (AIO) allows a program to initiate an I/O operation then continue with other work in parallel to the I/O operation. Oracle Database 11*g* often requires multiple server and user processes running at the same time. Therefore Oracle Database 11*g* takes full advantage of AIO services provided by AIX. AIO is implemented with AIO server processes. The configuration values of: minservers, maxservers and maxreqs control the AIO server configuration of AIX. The *Tuning IBM AIX 5L for an Oracle Database* whitepaper [1] has an "Asynchronous I/O" section that describes recommendations for the configuration values. GPFS configuration recommendations are also supplied. There is also a "Using Asynchronous I/O" section in the *Oracle Architecture and Tuning on AIX* whitepaper [2].

In AIX 5.3 AIO is disabled by default. However, in AIX 6.1 AIO is enabled by default. When upgrading to AIX 6.1, the AIO setting will not be changed.

Note, if you use the AIX filemon command, the AIO API calls of aio_read, aio_write and lio_listio are not included in the filemon report.

Concurrent I/O and direct I/O

There is file system I/O tuning information in the *Tuning IBM AIX 5L for an Oracle Database* whitepaper [1], the *Oracle Architecture and Tuning on AIX* whitepaper [2] and the *Direct I/O or Concurrent I/O on AIX 5L* My Oracle Support note [6]. For Oracle Database 11*g*, the database defaults to asynchronous I/O (AIO) enabled and concurrent I/O (CIO) disabled. In general, a good starting point is to set the filesystemio_options=setall, in your init*.ora configuration file. This setting will enable AIO (which is the default) and CIO operation. CIO operation is built upon direct I/O (DIO) with the additional function of inode locking. Note, there may be workloads (eg. sequential reads) where cached I/O performs better than CIO.

When using CIO/DIO, the Oracle setting of DB_FILE_MULTIBLOCK_READ_COUNT (the maximum number of blocks read in one I/O operation during a sequential scan) needs to be considered. Also, the alignment of the database blocksize and the file system block size (agblksize) has to be considered. These two topics are addressed by the documents in the previous paragraph. They are also addressed in the *Oracle 9i & 10g on AIX 5L: Tips and Considerations* whitepaper [7].

When not using CIO/DIO, look at the suggested settings in the "AIX sequential read ahead" section of the *Oracle Architecture and Tuning on AIX* whitepaper [2].

Note, some third-party utilities and file system utilities (eg. cp, dd and cpio) are not CIO/DIO aware. In these situations, if the database is started with filesystemio_options=setall, you could shutdown the database, thereby ensuring files are not open with CIO/DIO and proceed with the utilities using cached I/O. Or you could take advantage of the cio/dio mount options, specify filesystemio_options=asynch (the default) and the database and utilities can concurrently use CIO/DIO. If the CIO/DIO mount options are

used, file system read ahead will also be disabled. The utilities may need to have their default I/O transfer size increased to maintain acceptable throughput. This topic is also addressed in the *Oracle 9i & 10g on AIX 5L: Tips and Considerations* whitepaper [7].

File system cache size

In AIX 5.3 the default Virtual Memory Manager (VMM) settings are not optimal for a database environment. In AIX 6.1 the VMM defaults have been changed to be much more suitable for a database workload. When upgrading from AIX 5.3 to AIX 6.1, the VMM settings will not be changed. So, VMM settings need to be changed in the upgraded AIX 6.1 as needed for Oracle database. Refer the following section “Tuning resources” and the document “Tuning IBM AIX 5.3 and AIX 6.1 for Oracle Database (whitepaper) link.

For an Oracle database workload, we need to ensure the computational pages used for Oracle code, SGA and PGA remain resident in memory. The Oracle database buffer cache already provides caching of database files. Therefore, the file system cache size should be tuned (using the VMM settings) to favor computational pages over file pages. Check the “Memory and Paging” chapter in the *Oracle Architecture and Tuning on AIX* whitepaper [2] for recommended VMM settings. Note, these settings are a **suggested starting point**. If you have already tuned your system, do not revert to these VMM settings.

RAC IPC

Oracle RAC 11*g* uses the User Datagram Protocol (UDP) for interprocess communication (IPC) between nodes. The *Oracle Architecture and Tuning on AIX* whitepaper [2] indicates how to tune the UDP kernel settings in the “Network Tuning” section.

General Parallel File System

The *GPFS Concepts, Planning and Installation Guide* [3] contains a “GPFS use with Oracle” chapter that contains Oracle configuration and tuning considerations. My Oracle Support note 302806.1 [4] contains a “GPFS Tuning Requirements for Oracle” section which describes the latest tuning information.

Oracle process memory footprint

The AIXTHREAD_SCOPE environment variable can be used for control if an AIX process runs with process-wide contention scope (the default) or with system-wide contention scope. System-wide contention scope significantly reduces the memory required for each database process. AIX operates most effectively with Oracle Database 11*g* and Oracle RAC when using system-wide contention scope (AIXTHREAD_SCOPE=S). See the “Tuning Memory” chapter of the *Tuning IBM AIX 5L for an Oracle Database* whitepaper [1] for a detailed description of the AIXTHREAD_SCOPE parameter.

Tuning resources

These are the documents referenced throughout this section.

1. Tuning IBM AIX 5.3 and AIX 6.1 for Oracle Database (whitepaper), http://www.ibm.com/developerworks/forums/servlet/JiveServlet/download/747-332466-14479945-364618/Oracle_tuning_AIX_Jan09.pdf

2. Oracle RAC on IBM AIX best practices in memory tuning and configuring for system stability http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101513
3. Oracle Architecture and Tuning on AIX (whitepaper), http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100883
4. GPFS: Concepts, Planning and Installation Guide (GPFS document), http://publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp?topic=/com.ibm.cluster.gpfs.doc/gpfsbooks.html
5. IBM GPFS and Oracle RAC on AIX 5L and IBM pSeries (My Oracle Support note 302806.1), https://support.oracle.com/CSP/ui/flash.html.
6. GPFS FAQs available here, http://www-03.ibm.com/systems/clusters/software/gpfs.html
7. Direct I/O or Concurrent I/O on AIX 5L (My Oracle Support note 272520.1), https://support.oracle.com/CSP/ui/flash.html.
8. Oracle 9*i* & 10*g* on IBM AIX5L: Tips and Considerations (whitepaper), http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100556

These are supplemental tuning resources.

- AIX 6 Performance Management (AIX documentation), http://publib.boulder.ibm.com/infocenter/systems/scope/aix/topic/com.ibm.aix.doc/doc/base/performance.htm
- AIX 5L Performance Management Guide (AIX documentation), http://publib.boulder.ibm.com/infocenter/pseries/v5r3/topic/com.ibm.aix.doc/doc/base/performance.htm
- AIX 5L Practical Performance Tools and Tuning Guide (Redbook), http://www.redbooks.ibm.com/abstracts/sg246478.html?Open
- Performance Management for System p, http://www-03.ibm.com/systems/p/support/pm/index.html
- Configuring IBM TotalStorage for Oracle OLTP Applications (whitepaper), http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100319
- Configuring IBM System Storage DS4000 Series for Oracle Database Applications (whitepaper), http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100780
- Diagnosing Oracle Database Performance on AIX Using IBM NMON and Oracle Statspack Reports (whitepaper), http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100880
- Improving Database Performance with AIX Concurrent I/O (whitepaper), http://www.ibm.com/servers/aix/whitepapers/db_perf_aix.pdf
- Diagnosing Oracle Database Performance on AIX Using IBM NMON and Oracle Statspack Reports (whitepaper), http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101720
- RAC Starter Kit and Best Practices, “My Oracle Support” note 811293.1
- Additional information about Oracle10g and 11g database compatibility with AIX, refer “My Oracle Support” note 889464.1

Summary

This document gathers together the key technical topics that need to be considered in planning or implementing Oracle Database 11*g* R2 or Oracle RAC 11*g* R2 with AIX. In almost all cases, there are reinforcing references included in each section. The resources appendix below contains pointers to general documentation and additional supporting documents.

Appendix 1: Resources

These Web sites and documents provide useful references to supplement the information contained in this document:

- Oracle Database 11*g* Documentation Library,
  http://www.oracle.com/pls/db111/homepage

- Oracle Clusterware and RAC documentation,
  http://www.oracle.com/pls/db111/portal.portal_db?selected=16&frame=

- Oracle RAC SIG,
  http://www.oracleracsig.org

- IBM and Oracle Support Process,
  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101878

- Raw Devices and Cluster Filesystems with RAC, My Oracle Support note 183408.1,
  https://support.oracle.com/CSP/ui/flash.html.

- Oracle My Oracle Support note 341507.1: Oracle Products on Linux on IBM POWER,
  https://support.oracle.com/CSP/ui/flash.html.

- Oracle and IBM System Storage
  http://www-03.ibm.com/systems/storage/solutions/compatibility/oracle/index.html

About the author

Ravisankar Shanmugam is an IT Specialist with the IBM Oracle International Competency Center (ICC) based in San Mateo, California. He writes technical sales collateral and sizing information for Oracle products on AIX. He also provides support for AIX and Power Systems projects at the IBM Oracle ICC and onsite at Oracle Headquarters in Redwood Shores and Pleasanton, CA.

Trademarks and special notices

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

IBM, the IBM logo, AIX, AIX 5L, AS/400, eServer, GPFS, PowerHA, iSeries, Micro-Partitioning, Power, PowerHA, PowerVM, Power Architecture, POWER, POWER3, POWER4, POWER5, POWER6, POWER7, PowerPC, Power Systems, pSeries, RS6000, System i, System i5, System p, and System p5 are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

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

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

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

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

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

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local IBM office or IBM authorized reseller for the full text of the specific Statement of Direction.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in IBM product announcements. The information is presented here to communicate IBM's current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.