Advanced Technical Skills

AIX Performance:
Configuration & Tuning for Oracle & Oracle RAC

Vijay Adik
vadik@us.ibm.com
ATS - Oracle Solutions Team

December 5, 2013

Legal information

The information in this presentation is provided by IBM on an "AS IS" basis without any warranty, guarantee or assurance of any kind. IBM also does not provide any warranty, guarantee or assurance that the information in this paper is free from errors or omissions. Information is believed to be accurate as of the date of publication. You should check with the appropriate vendor to obtain current product information.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

IBM, RS6000, System p, AIX, AIX 5L, GPFS, and Enterprise Storage Server (ESS) are trademarks or registered trademarks of the International Business Machines Corporation.

Oracle, Oracle9i, Oracle10g and Oracle 11g are trademarks or registered trademarks of Oracle Corporation.

All other products or company names are used for identification purposes only, and may be trademarks of their respective owners.

December 5, 2013

Agenda

- **AIX Configuration Best Practices for Oracle**
  - **Memory**
  - **CPU**
  - **I/O**
  - **Network**
  - **Miscellaneous**

December 5, 2013

AIX Configuration Best Practices for Oracle

- **The suggestions presented here are considered to be basic configuration "starting points" for general Oracle workloads**

- **Your workloads may vary**

- **Ongoing performance monitoring and tuning is recommended to ensure that the configuration is optimal for the particular workload characteristics**

December 5, 2013

Agenda

- **AIX Configuration Best Practices for Oracle**

  – **Memory**

  – CPU

  – I/O

  – Network

  – Miscellaneous

Virtual Memory Manager (VMM) Tuning

- **The AIX "vmo" command provides for the display and/or update of several parameters which influence the way AIX manages physical memory**

  – **The "-a" option displays current parameter settings**

    ➢ **vmo –a**

  – **The "-o" option is used to change parameter values**

    ➢ **vmo –o minfree=1440**

  – **The "-p" option is used to make changes persist across a reboot**

    ➢ **vmo –p –o minfree=1440**

> **On AIX 5.3, number of the default "vmo" settings are not optimized for database workloads and should be modified for Oracle environments**

December 5, 2013

Kernel Parameter Tuning – AIX 6.1/7.1

- **AIX 6.1/7.1 configured by default to be 'correct' for most database workloads.**

- **Many tunable are classified as 'Restricted':**

  - Only change if AIX Support says so

  - Parameters will not be displayed unless the '-F' option is used for commands like vmo, no, ioo, etc.

- **When migrating from AIX 5.3 to 6.1/7.1, parameter override settings in AIX 5.3 will be transferred to AIX 6.1/7.1 environment**

December 5, 2013

General Memory Tuning

- **Two primary categories of memory pages: Computational and File System**

- **AIX will always try to utilize all of the physical memory available (subject to vmo parameter settings)**
  - ✓ What is not required to support current computational page demand will tend to be used for filesystem cache
  - ✓ Raw Devices and filesystems mounted (or individual files opened) in DIO/CIO mode do not use filesystem cache

- **Requests for new memory pages are satisfied from the free page list**
  - ✓ Small reserve of free pages maintained by "stealing" Computational or File pages
  - ✓ AIX uses "demand paging" algorithm – generally not written to swap file until "stolen"

December 5, 2013
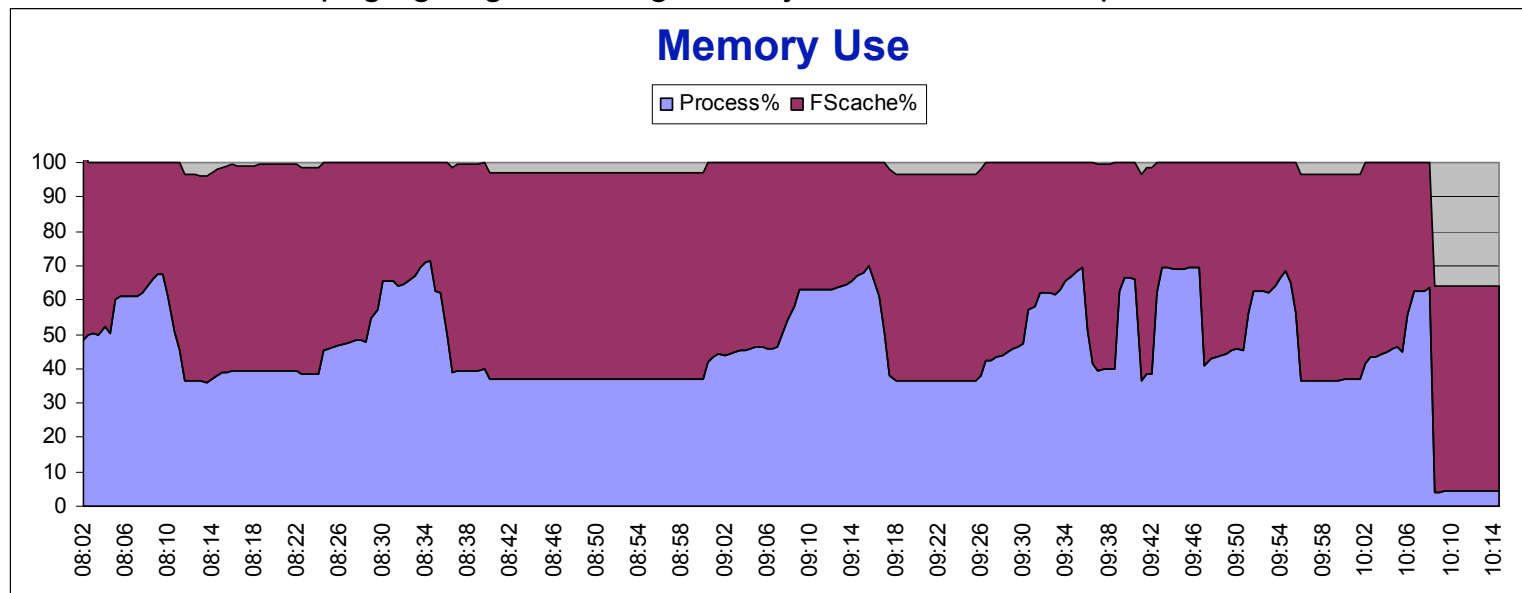
Memory Usage in an Oracle Database Environment

Computational

- Some used for AIX kernel processing
- Some used by Oracle/client executable programs
- Includes Oracle SGA and PGA memory

File System Cache

- May be used for caching or pre-fetching of Oracle .dbf files
  - Only for local file system based (non-RAC) environments where Direct I/O (or Concurrent I/O) is not used
- May be used for other Oracle related files
  - Archive logs, export/import files, backups, binaries, etc.
- May be used for non-Oracle related files
  - Application files, system files, etc.

Virtual Memory Management Priorities

- Always want to keep computational pages in memory -- System paging/swapping may degrade Oracle/application performance
  - Allocate enough physical memory to support computational footprint requirement + small file cache
  - When necessary, steal filesystem pages, not computational

September 2012
December 5, 2013

AIX System Paging Concepts & Requirements

- **By default, AIX uses a "demand paging" policy**

  - For Oracle DB, the goal is ZERO system paging activity

  - Filesystem pages written back to filesystem disk (if dirty); never to system paging space

  - Unless otherwise specified, computational pages are not written to paging space unless/until they are stolen by lrud. (*1)

- **Once written to paging space, pages are not removed from paging space until the process associated with those pages terminates**

  - For long running processes (e.g. Oracle DB), even low levels of system paging can result in significant growth in paging space usage over time

  - Paging space should be considered a fail-safe mechanism for providing sufficient time to identify and correct paging issues, not a license to allow ongoing system paging activity

- **Paging space allocation Rule-of-Thumb:**

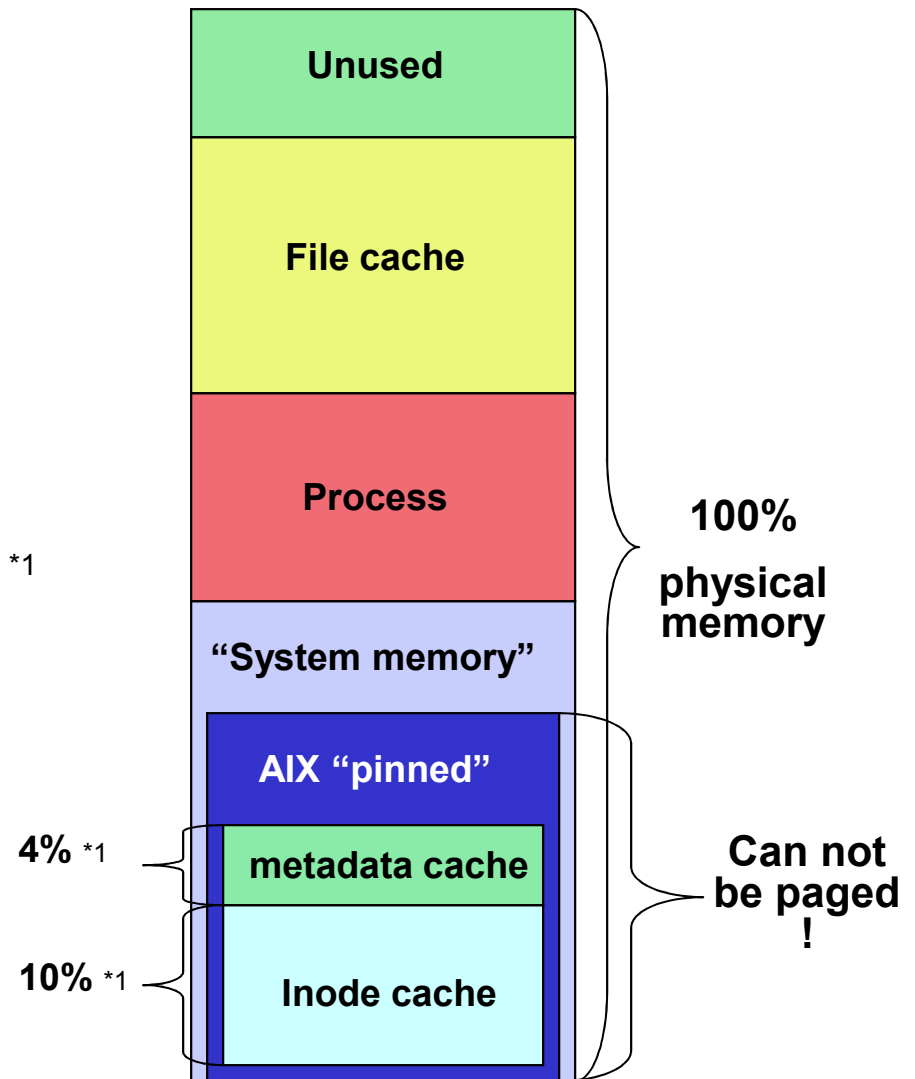 - ½ the physical memory + 4 GB, with the following cap:

Resolve paging issues quickly:

- Reduce effective minimum file system cache size (minperm)
- Reduce Oracle SGA or PGA size
- Add physical memory

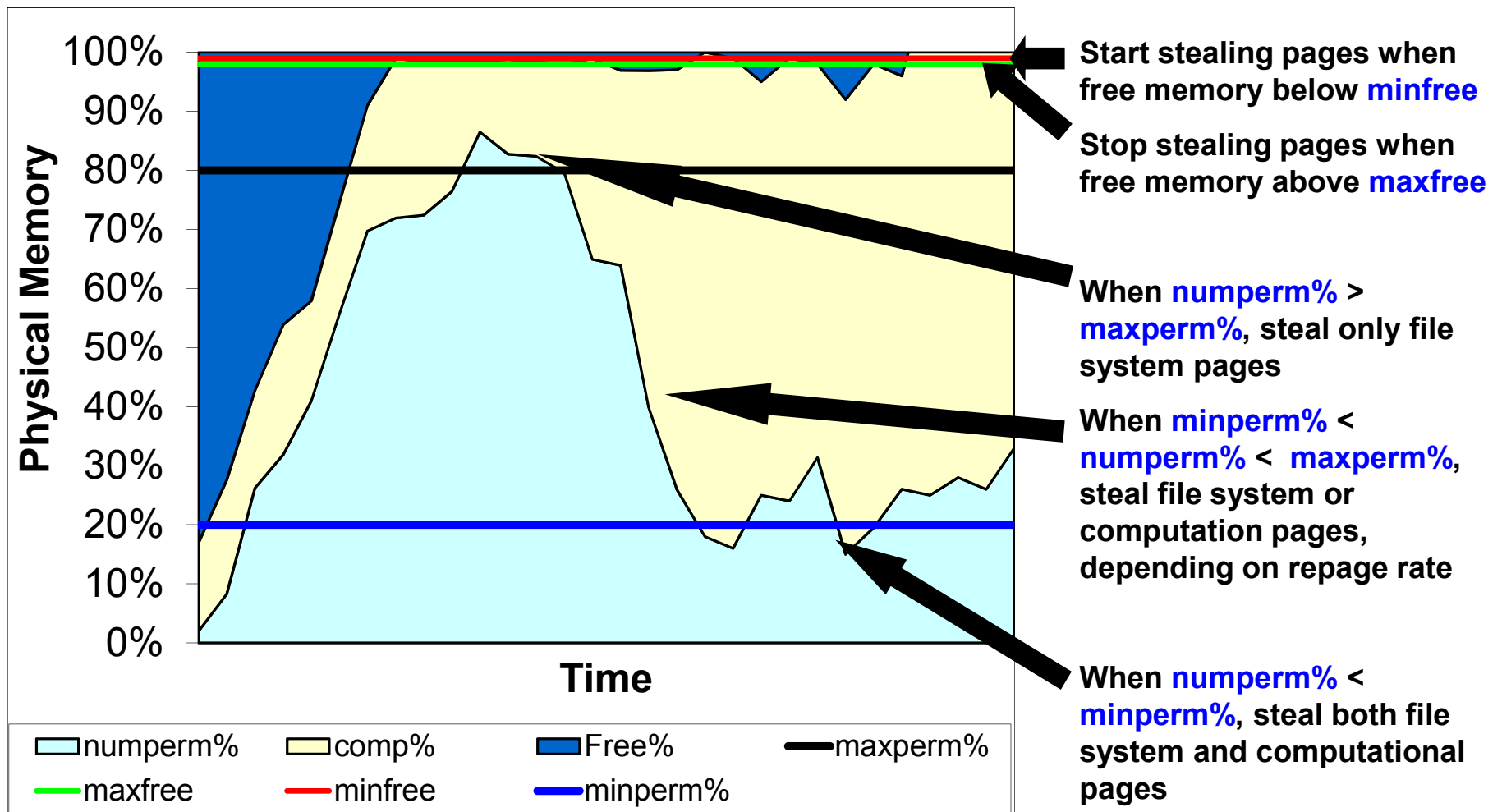| Physical Memory lower or equal to	Paging Space Max
128GB	60GB
256GB	100GB
512GB	150GB
1TB	200GB

December 5, 2013

JFS2 inode / metadata caches

- **JFS2 utilizes two caches - one for inodes and one for metadata**

- **Caches grow in size until maximum size is reached before cache slots are reused**

- **Default values are tuned for a file server!**

- **Each entry in the inode cache requires about 1KB of physical memory**

  → 1MB of memory can cache about 1000 files

- **Configured via ioo parameters:**

  – j2_inodeCacheSize   (Default: 400 = 10%) *1

  – j2_metadataCacheSize (Default: 400 = 4%) *1

- **The current memory use can be verified via:**
  cat /proc/sys/fs/jfs2/memory_usage
      metadata cache: 31186944
      inode cache: 34209792
      total: 65396736

*1 Note: Default values in AIX 7.1 are 200 (5%) , 200 (2%)

©2013 IBM Corporation

December 5, 2013

Virtual Memory Management (VMM) Thresholds

Start stealing pages when free memory below minfree

Stop stealing pages when free memory above maxfree

When numperm% > maxperm%, steal only file system pages

When minperm% < numperm% < maxperm%, steal file system or computation pages, depending on repage rate

When numperm% < minperm%, steal both file system and computational pages

Legend: numperm%   comp%   Free%   maxperm%   maxfree   minfree   minperm%

Page Steal Method

- **Historically, AIX maintained a single LRU list which contains both computational and filesystem pages.**

  - In environments with lots of computational pages that you want to keep in memory, LRUD may have to spend a lot of time scanning the LRU list to find an eligible filesystem page to steal

- **AIX 6.1 introduced the ability to maintain separate LRU lists for computational vs. filesystem pages.**

  - Also backported to AIX 5.3

- **New page_steal_method parameter**

  - Enabled (1) by default in 6.1/7.1, disabled (0) by default in 5.3
  - Requires a reboot to change
  - Recommended for Oracle DB environments

December 5, 2013

Large Segment Aliasing (AIX 6.1 TL06, AIX 7.1)

- **Feature allows user applications to "automagically" use 1TB segments.**

  – 1 SLB entry in POWER7 can now address 1TB of memory.

  - Segment Lookaside Buffer (SLB) fault issue no longer relevant
  - Immediate performance boost for applications, new and legacy

- **Significant changes under the covers**

  – New address space allocation policy

  - Attempts to group address space requests together to facilitate 1TB aliasing.

  – Once certain allocation size thresholds have been reached, OS automatically aliases memory with 1TB aliases.

  - 256MB segments still exist for handling IO

- **Aliasing only available for shared memory regions at this point.**

- vmo –p –o esid_allocator = 1   and shm_1tb_unsh_enable = 0

Recommended vmo "Starting Points" - Review

Parameter	Recommend Value	AIX 7.1 Default	AIX 7.1 Restricted	AIX 6.1 Default	AIX 6.1 Restricted	AIX 5.3 Default
minperm%	3	3	No	3	No	20
maxperm%	90	90	Yes	90	Yes	80
maxclient%	90	90	Yes	90	Yes	80
strict_maxclient	1	1	Yes	1	Yes	1
strict_maxperm	0	0	Yes	0	Yes	0
lru_file_repage	0	N/A	N/A	0	Yes	1 or 0(*1)
lru_poll_interval	10	10	Yes	10	Yes	10
minfree	960	960	No	960	No	960
maxfree	1088(*2)	1088	No	1088	No	1088
page_steal_method	1	1	Yes	1	Yes	0
memory_affinity	1	1	Yes	1	Yes	1
v_pinshm	0	0	No	0	No	0
lgpg_regions	0	0	No	0	No	0
lgpg_size	0	0	No	0	No	0
maxpin%	Leave at Default	90(*3)	No	80(*3)	No	80

*1 Depending on AIX 5.3 TL level   *2 Do not reduce below default *3 Depends on LSA use – LSA active → 90, otherwise 80

December 5, 2013

Understanding Memory Pools

- **Memory cards are associated with every Multi Chip Module (MCM), Dual Core Module (DCM) or Quad Core Module (QCM) in the server**

  – The Hypervisor assigns physical CPUs to a dedicated CPU LPAR (or shared processor pool) from one or more MCMs, DCMs or DCMs

  – For a given LPAR, there will normally be at least 1 memory pool for each MCM, DCM or QCM that has contributed

  processors to that LPAR or shared processor pool

- **By default, memory for a process is allocated from memory associated with the processor that caused the page fault.**

- **Memory pool configuration is influenced by the VMO parameter "memory_affinity"**

  – Memory_affinity=1   means configure memory pools based on physical hardware configuration (DEFAULT)

  – Memory_affinity=0 means configure roughly uniform memory pools from any physical location

- **Number can be seen with 'vmstat –v |grep pools'**

- **Size can only be seen using KDB**

- **LRUD operates per memory pool**

December 5, 2013
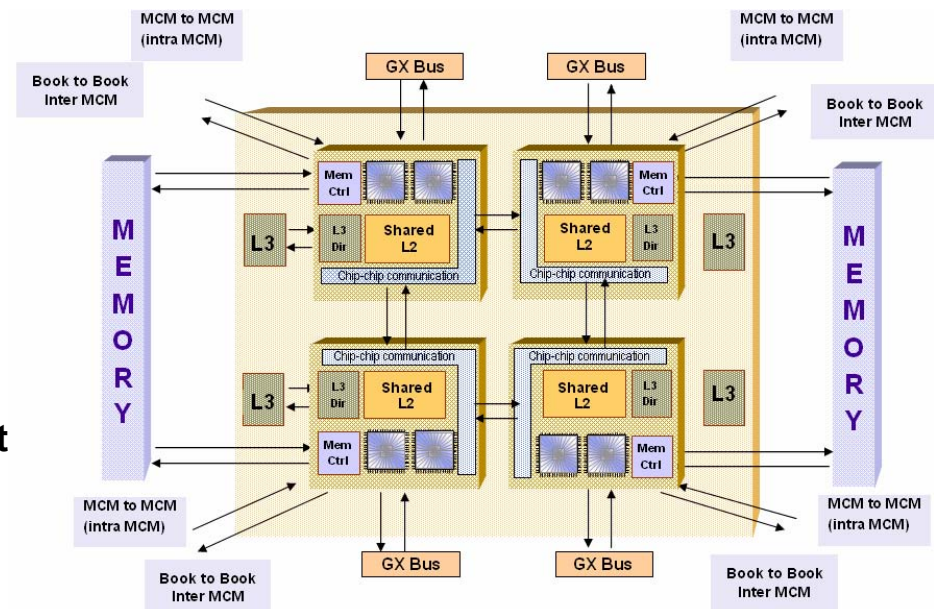
Oracle Memory and Memory Affinity

- **Oracle SGA is "striped" across all the available memory in the LPAR**
  - If the LPAR configuration has a combination of near, local and far memory allocated to it, SGA will be (more or less) evenly spread across all of it
  - The greater the number of CECs involved, the greater the likelihood of remote memory accesses

- **Oracle PGA for a given process tends to be allocated in the near memory of the processor that process was running on when the memory was allocated**

  - The AIX dispatcher will attempt to maintain affinity between a given process and the processor that process gets scheduled on
  - *rsets* may (optionally) be used to force affinity to a subset of available processors (e.g. those on a given chip, or within a given CEC), although this could potentially cause dispatching delays in heavily loaded environments

- **"vmo" enhanced_affinity_private** ("vmo" restricted parameter)

  - The percentage of application data that is to be allocated local, with the remaining memory to be striped across all available memory in the LPAR
  - Default value is 20% in AIX 6.1 TL5 and 40% in AIX 6.1 TL6+ and AIX 7.1
  - A value of 100% has been used in a number of Oracle performance benchmarks, but before changing a restricted parameter, contact IBM support.

September 2012
December 5, 2013

AIX Multiple Page Size Support

- **4K (Default)**
  - **All memory pages in the filesystem cache are 4K in size!**
  - **Can be paged to paging space**
  - **4K pages can be combined to a 64K page if not enough 64K pages are available**

- **64K, available with POWER5+ and later & AIX 5.3 TL4+**
  - **Can be paged to paging space**
  - **Can be converted to 4K pages if not enough 4K pages are available**
  - **Kernel page size used in AIX 5.3 TL4+ and above (can be configured)**
  - **Can be utilized for application code, data and stack as well, but requires specific configuration**

- **16M available with POWER4 hardware (or later) (also referred to as Large Pages)**
  - **Requires pinned memory and explicit configuration**
  - **Can not be paged to paging space**

- **16G available with POWER5+ & AIX 5.3 TL4+**
  - **Cannot be used with Oracle**
  - **Can not be paged to paging space**

December 5, 2013

4K - 64K - 16MB Page Dynamics

December 5, 2013
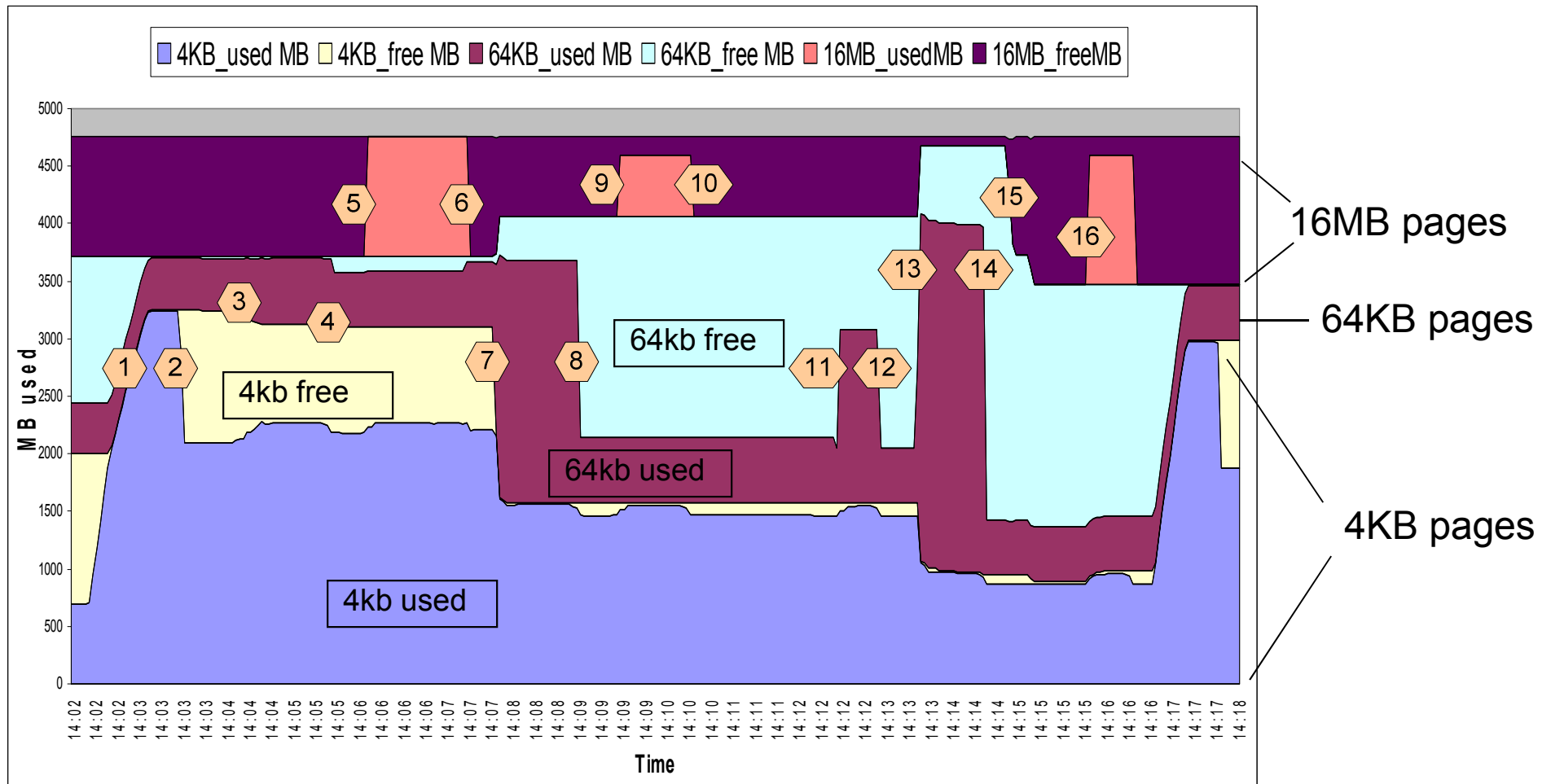
AIX Multiple Page Size Support

- **User/Application must request preferred page size**

    – 64K page size is very promising, since they do not need to be configured/reserved in advance or pinned

    - export LDR_CNTRL=DATAPSIZE=64K@TEXTPSIZE=64K@STACKPSIZE=64K@*SHMPSIZE=64K* to use the 64K pagesize for stack, data & text

    – Will require Oracle to explicitly request the page size (10.2.0.4 & up plus Oracle patch# 7226548)

    – If preferred size not available, the largest available smaller size will be used

    - Current Oracle versions will end up using 64KB pages even if SGA is not pinned

- **Refer: http://www-03.ibm.com/systems/resources/systems_p_os_aix_whitepapers_multiple_page.pdf**

December 5, 2013

Large Page Support 16mb (optional)

Pinning shared memory

- AIX Parameters
  - vmo –p –o v_pinshm = 1
  - Leave maxpin% at the default of 80% unless the SGA exceeds 77% of real memory
    - Vmo –p –o maxpin%=[(total mem-SGA size)*100/total mem] + 3
- Oracle Parameters
  - LOCK_SGA = TRUE

Enabling Large Page Support

- vmo –p –o lgpg_size = 16777216 –o lgpg_regions=(SGA size / 16 MB)

Allowing user oracle to use Large Pages

- chuser capabilities=CAP_NUMA_ATTACH,CAP_BYPASS_RAC_VMM,CAP_PROPAGATE oracle

Using Monitoring Tools

- svmon –G
- svmon –P

Oracle metalink note# 372157.1

Note: It is recommended not to pin SGA, as long as you had configured the VMM, SGA & PGA properly.

December 5, 2013

Determining SGA size

SGA Memory Summary for DB: test01 Instance: test01 Snaps: 1046 -1047

SGA regions	Size in Bytes
-------------------------------	----------------
Database Buffers	16,928,210,944
Fixed Size	768,448
Redo Buffers	2,371,584
Variable Size	1,241,513,984
	----------------
sum	18,172,864,960

lgpg_regions = 18,172,864,960 / 16,777,216 = 1084 (rounded up)

Oracle Memory Structures Allocation

- ## 11g : Automatic Memory Management (AMM)

  - memory_target (dynamic parameter) – specifies the total memory size to be used by the instance SGA and PGA. Exchanges between SGA and PGA are done according to workload requirements

  - If sga_target and pga_aggregate_target are not set, the policy is to give 60% of memory_target to the SGA and 40% to the PGA.

  - memory_max_target (static parameter) – specifies the maximum memory size for the database instance.

  - *To use Automatic Memory Management, memory_target must be >0.*

  - *Can not be used together with "LOCK_SGA=TRUE*

- *See Metalink notes 443746.1 and 452512.1 explaining AMM and these new parameters.*

AMM dynamic resizing of the shared pool can cause a fair amount of "cursor: pin s" wait time. One strategy to minimize this is to set minimum sizes for memory areas you particularly care about.

In addition you can change the frequency how often AMM analyzes and adjusts the memory distribution. See: Metalink note: 742599.1 (_memory_broker_stat_interval)

December 5, 2013

SGA_MAX_SIZE and LOCK_SGA implications (11g, 10.2.4.0+)

- **LOCK_SGA=false   Preferred**
  - Oracle dynamically allocates memory for the SGA only as needed up to the size specified by SGA_TARGET
  - SGA_TARGET may be dynamically increased, up to SGA_MAX_SIZE
  - 64K pages automatically used for SGA if supported in the environment. If needed, 4K (or 16M) pages are converted to 64K pages.

- **LOCK_SGA=true   Discouraged**
  - Oracle Automatic Memory Management (AMM) cannot be used (MEMORY_TARGET)
  - Oracle pre-allocates all memory as specified by SGA_MAX_SIZE and pins it in memory, even if it's not all used (i.e. SGA_TARGET < SGA_MAX_SIZE)
  - If sufficient 16M pages are available those will be used. Otherwise, all the SGA memory will be allocated from 64K (if supported) or 4K pages (if 64K pages are not supported). If needed, 4K (or 16M pages will be converted to 64K pages, but 16M pages are never automatically created.
  - If a value for sga_max_size is specified larger than the amount of available memory for computational pages, the system can become unresponsive due to system paging.
  - If the specified SGA_MAX_SIZE is much larger than the currently available pages on the combined 64K and 16M page free lists, the database startup can fail with error: *"IBM AIX RISC System/6000 Error: 12: Not enough space"*. In this case re-try to start the database.

December 5, 2013

AIX dynamic LPAR with Oracle AMM

Initial configuration

Memory_max_size
= 18 GB

Real memory
= 12 GB

AIX
+ free

Memory_target
= 8 GB

SGA
+
PGA

Scenario :

Oracle tuning advisor indicates that
SGA+PGA need to be increased to 11GB:

memory_target can be increased
dynamically to 11GB but real memory is only
12GB, so it needs to be increased as well.

▶ Step 1 - Increase physical memory
allocated to the LPAR to 15 GB

ssh hscroot@hmc "chhwres -r mem
-m <system> -o a -p <LPAR name> -q 3072"

▶ Step 2 - Increase SGA+PGA allocated
to the instance of the database to 11GB

alter system set memory_target=11GB;

Final configuration

Memory_max_size
= 18 GB

Real memory
= 15 GB

AIX
+ free

Memory_target
= 11 GB

SGA
+
PGA

➔ Memory allocated to the system has been increased dynamically, using AIX DLPAR
➔ Memory allocated to Oracle (SGA and PGA) has been increased on the fly

December 5, 2013

Memory Usage in an Oracle Environment

Computational
- Some used for AIX kernel processing
- Some used by Oracle/client executable programs
- Includes Oracle SGA and PGA memory

File System Cache
- May be used for caching or prefetching of Oracle .dbf files
  - Only for local filesystem based (non-RAC) environments where Direct I/O (or Concurrent I/O) is not used
- May be used for other Oracle related files
  - Archive logs, export/import files, backups, binaries, etc.
- May be used for non-Oracle related files
  - Application files, system files, etc.

Virtual Memory Management Priorities
- Always want to keep computational pages in memory -- System paging/swapping may degrade Oracle/application performance
  - Allocate enough physical memory to support computational footprint requirement + small file cache
  - When necessary, steal filesystem pages, not computational

ACTION

1 - System Startup

2 - DB Activity (needs more memory)

3 - => Paging Activity

MEMORY | PAGING SPACE

Performance degradation

FS CACHE

SGA + PGA

Programs

Kernel

SGA + PGA

Oracle Server Architecture – Memory Structures

Figure 8-1 Oracle Database Memory Structures
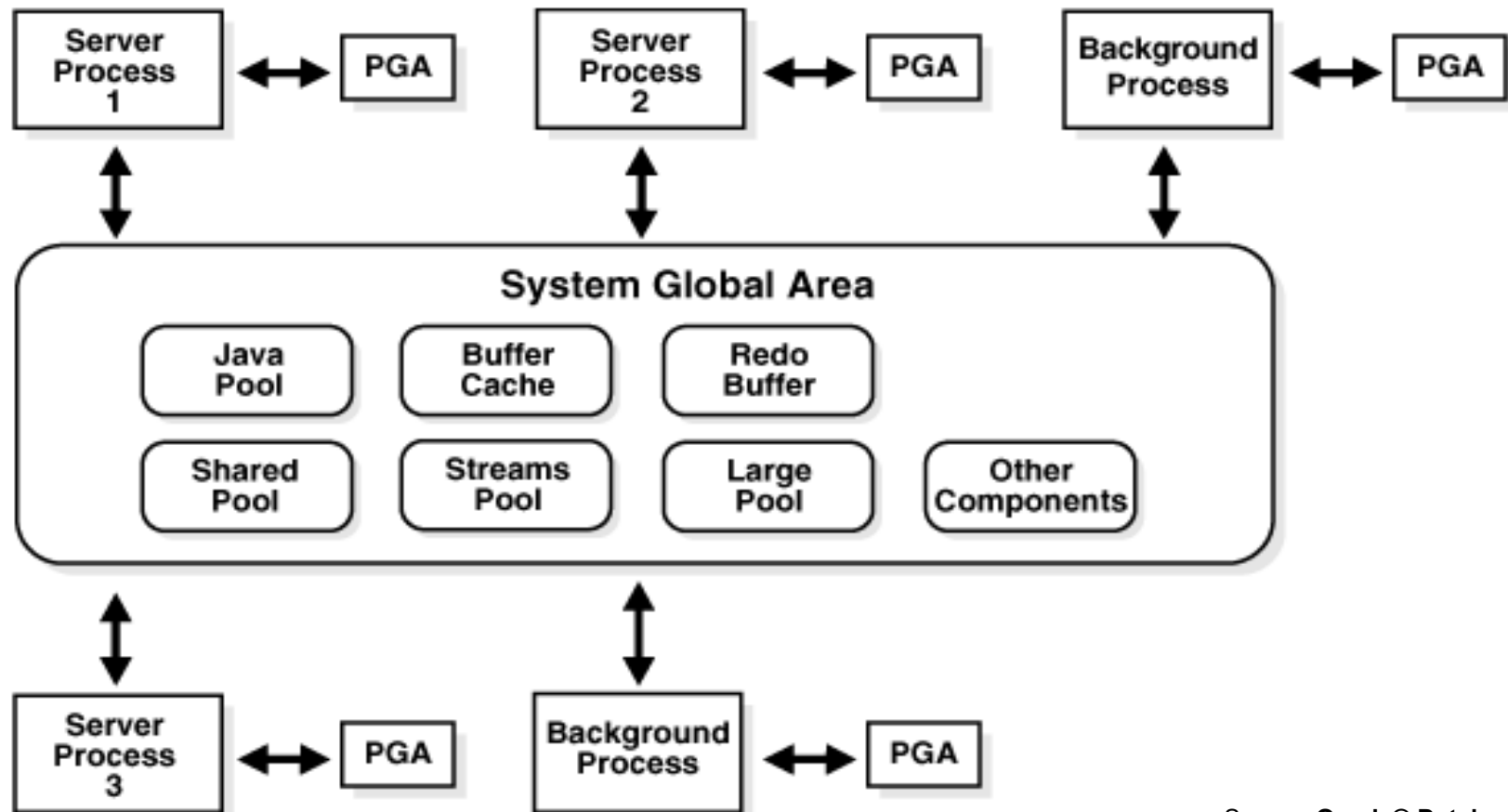
Source: **Oracle® Database Concepts**
11g Release 1 (11.1)
Part Number B28318-05

December 5, 2013

Displaying Memory Usage Statistics

The 'vmstat' command provides information on current memory usage: (reported in 4K pages)

vmstat -v

```
  1048576  memory pages
  1002006  lruable pages
   812111  free pages
        1  memory pools
   141103  pinned pages
     80.0  maxpin percentage
      3.0  minperm percentage
     90.0  maxperm percentage
      3.2  numperm percentage
    32779  file pages
      0.0  compressed percentage
        0  compressed pages
      0.0  numclient percentage
     90.0  maxclient percentage
        0  client pages
        0  remote pageouts scheduled
        0  pending disk I/Os blocked with no pbuf
        0  paging space I/Os blocked with no psbuf
     2484  filesystem I/Os blocked with no fsbuf
        0  client filesystem I/Os blocked with no fsbuf
        0  external pager filesystem I/Os blocked with no fsbuf
```

December 5, 2013

Displaying Memory Usage Statistics

The 'svmon -G' command provides information on current memory usage per page size: (general numbers are reported in 4K pages)

```
# "svmon –G" on AIX 5.3 / "vmstat –P ALL " on AIX 6.1 & 7.1
```

	size	inuse	free	pin	virtual
memory	1179648	926225	290287	493246	262007
pg space	1572864	5215			

	work	pers	clnt	other
pin	91390	0	0	74176
in use	258573	4316	335656	

PageSize		PoolSize	inuse	pgsp	pin	virtual
s	4 KB	–	477713	5215	94606	141175
m	64 KB	–	7552	0	4435	7552
L	16 MB	80	0	0	80	0

December 5, 2013

AIX Paging Space

Allocate Paging Space:

- Configure Server/LPAR with enough physical memory to satisfy memory requirements
- With AIX demand paging, paging space does not have to be large
  - **Provides safety net to prevent system crashes when memory overcommitted.**
- Generally, keep within internal drive or high performing SAN storage

Monitor paging activity:

- vmstat -s
- sar -r
- nmon

Resolve paging issues:

- Reduce file system cache size (MAXPERM, MAXCLIENT)
- Reduce Oracle SGA or PGA (9i or later) size
- Add physical memory

Do not over commit real memory!

December 5, 2013

A few important notes

AIX 6.1 kernel pinning :

- vmo -r -o vmm_klock_mode=2
  - ➢ It is enabled by default in 7.1 and can be enabled on 6.1 TL6.
- There were cases of Oracle RAC node eviction due to higher memory allocation and/or bug/feature
- Make sure you are in this release : AIX 6.1 TL6 SP5+, plus APAR IZ95744

USLA heap issue:

- With the release of 11*g*R2, Oracle introduced a new feature called "on line patching" and is supported on AIX 6.1 TL02 and 7.1 TL0 onwards
- Upgrade to AIX 6.1 TL07 *or* AIX 7.1 TL01, and then apply the Oracle bug patch 13443029 for database 11.2.0.2 and 11.2.0.3
- If you cannot upgrade to AIX 6.1 TL07 or 7.1 TL01, then apply Oracle patch 10190759, which will have a reduced online-patch functionality
- It is recommended to use AIX 6.1 TL07 SP02 or AIX 7.1 TL01 SP02

P7 special patch:

- Patch 6784747 - Wasted memory in koh-kghu allocations. Free extents of memory might not get used.
- This bug impacts 10.2.0.4/5 and 11.1.0.x releases. This fix is not included in any of the PSU. This issue is fixed in Oracle 11.2.0.1 onwards

Agenda

- **AIX Configuration Best Practices for Oracle**

  – Memory

  – **CPU**

  – I/O

  – Network

  – Miscellaneous

December 5, 2013

CPU Considerations

Oracle Parameters based on the # of CPUs

- DB_WRITER_PROCESSES

- Degree of Parallelism

  - user level

  - table level

  - query level

  - MAX_PARALLEL_SERVERS or AUTOMATIC_PARALLEL_TUNING (CPU_COUNT * PARALLEL_THREADS_PER_CPU)

- CPU_COUNT

- FAST_START_PARALLEL_ROLLBACK – should be using UNDO instead

- CBO – execution plan may be affected; check explain plan

Physical, Logical, Virtual Layers

AIX 5.2 LPAR SMT=off	AIX 5.3 LPAR SMT=on	AIX 5.3 Micro Partition SMT=on	AIX 6.1 Micro Partition SMT=off	AIX 7.1 Micro Partition SMT-4	
	L L	L L L L L L L L		L L L L L L	**Logical**
				L L L L L L	
		V V V V	V V	V V V	**Virtual**
		2.1 Proc. Units	**0.8 Proc Units**	**1.2 Proc Units**	**Physical**
2 CPUs (dedicated)	**1 CPU (dedicated)**	Shared Pool 0		Pool 1	
		13 CPU Shared Processor Pool*			

16 CPU SMP Server
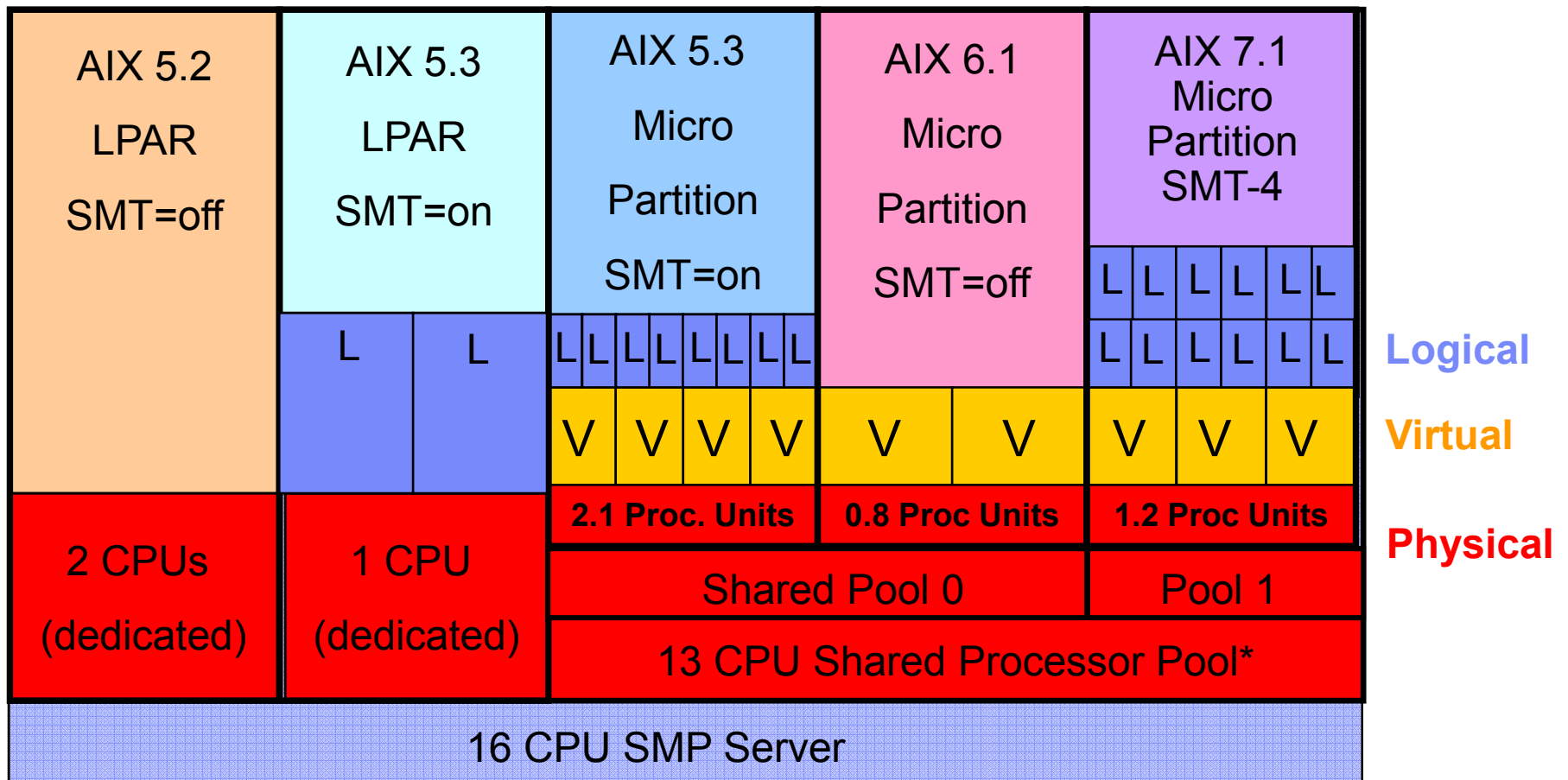
Think "PVL "   **P=Physical**   **V=Virtual**   **L=Logical (SMT)**

* All activated, non-dedicated CPUs are automatically placed into the shared processor pool.
   Only 2.1+0.8+1.2 = 4.1 processor units of "desired capacity" has been allocated from the pool of 13 CPUs

Virtual Shared Processor Pools - Benefits

Server with 12 processor cores

POWER6/7 Multiple shared pools:

- Can reduce the number of software licenses by putting a limit on the amount of processors an uncapped partition can use

- Up to 64 shared pools

LPAR4	LPAR5	LPAR6	LPAR7	LPAR8
Uncapped	Uncapped	Uncapped	Uncapped	Uncapped
AIX	AIX	AIX	AIX	AIX
		OAS	OAS	OAS
Oracle DB	Oracle DB	App 1	App2	App 3
VP = 5	VP = 4	VP = 4	VP = 6	VP = 3
Ent. = 2.5	Ent. = 1.70	Ent. = 2.00	Ent. = 2.00	Ent. = 1.00

LPAR1	LPAR2	LPAR3	Virtual Shared pool #1				Virtual Shared pool #2				
i5/OS	AIX	Linux	**Max Cap: 5 processors**				**Max Cap: 6 processors**				
	Oracle DB		Physical Shared Pool (9 processor cores)								
1	1	1	1	2	3	4	5	6	7	8	9

Oracle cores to license:
- 1 from dedicated partition **LPAR**2
- 5 from pool 1
- **= 6**

OAS cores to license:
- 6 from pool 2
- **= 6**

Oracle DB core – license factors:

Power5 and earlier:	**0.75**
Power6:	**1.0**
Power7:	**1.0**

December 5, 2013

Virtual Processor - Folding

- **Dynamically adjusting active Virtual Processors (VPs)**
  - System consolidates loads onto a minimal number of VPs
    - Scheduler computes utilization of VPs every second
      - If VPs needed to host physical utilization is less than the current active VP count, a VP is put to sleep
      - If VPs needed are greater than the current active VPs, more are enabled
  - On by default in AIX 5.3 ML3 and later
    - `vpm_xvcpus` **tunable**
    - `vpm_fold_policy` **tunable**

- **Increases processor utilization and affinity**
  - Inactive VPs don't get dispatched and waste physical CPU cycles
  - Fewer VPs can be more accurately dispatched to physical resources by the Hypervisor with potential for improved processor cache efficiency

- **When to adjust – Check with IBM support before changing!**
  - Burst/Batch workloads with short response-time requirements may need sub-second dispatch latency
    - Disable processor folding or manually tune the number of VPs
      - `# schedo -o vpm_xvcpus=[-1 | N]`
      - Where N specifies the number of VPs to enable in addition to the number of VPs needed to consume physical CPU utilization
      - A value of "-1" disables CPU folding

	SMT	DLPAR	Micro-Partition	LPM	WPAR	AME
AIX 5.2	⛔	✓	⛔	⛔	⛔	⛔
AIX 5.3	✓	✓	✓	✓	⛔	⛔
AIX 6.1	✓	✓	✓	✓	✓	✓
AIX 7.1	✓	✓	✓	✓	✓	✓
Oracle 9i	✓	✓	✓	⛔	⛔	⛔
Oracle 10g	✓	✓	✓	✓ *1	✓ *3	⛔
Oracle 11g	✓	✓	✓	✓ *2	✓ *4	✓ *5

***1** - 10.2.0.4 with minimum of AIX 5.3 TL8 SP4 or AIX 6.1 TL2 SP3
***2** - 11gR1 with AIX 6.1 TL3 SP1, 11gR2 with AIX 5.3 TL9 SP6, 11gR2 with AIX 6.1 TL4 SP1
***3** – 10.2.0.4 with minimum AIX 6.1 TL2 SP2 (No AIX 7.1 support at this time)
***4** – 11.2.0.2 with minimum AIX 6.1 TL2 SP2 (No AIX 7.1 support at this time)
***5** – 11.2.0.2 and above with minimum AIX 6.1 TL6 SP5 and AIX 7.1

Note: Live partition mobility is certified for 10.2.0.4 and AIX 5.3 TL8 SP4 & 6.1 TL2 SP3.

Oracle RAC 10.2.0.3 on VIOS 1.3.1.1 & AIX 5.3 TL07 and higher are certified.

Agenda

- **AIX Configuration Best Practices for Oracle**
  - Memory
  - CPU
  - **I/O**
  - Network
  - Miscellaneous

The AIX IO stack

			Application		▢ (cyan)
Raw disks	**Raw LVs**	**Logical file system**			
		JFS	JFS2	NFS	Other
			VMM		▢ (cyan)
		LVM (LVM device drivers)			
Multi-path IO driver (optional)					
Disk Device Drivers					▢ (cyan)
Adapter Device Drivers					▢ (cyan)
Disk subsystem (optional)					▢ (blue)
Disk					▢ (cyan)

▮ (blue) Write cache   ▢ (cyan) Read cache or memory area used for IO
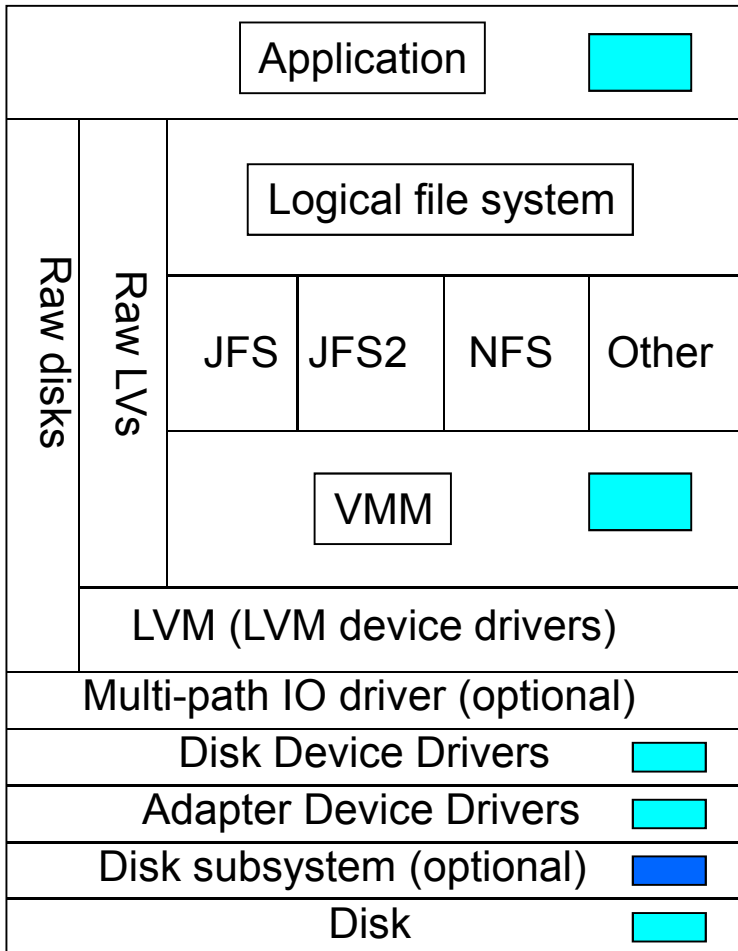
Application memory area caches data to
avoid IO

NFS caches file attributes
NFS has a cached filesystem for NFS clients

JFS and JFS2 cache use extra system RAM
JFS uses persistent pages for cache
JFS2 uses client pages for cache

Queues exist for both adapters and disks
Adapter device drivers use DMA for IO
Disk subsystems have read and write cache
Disks have memory to store commands/data

IOs can be coalesced (good) or split up (bad) as they go thru the IO stack

AIX Filesystems Mount options

- **Journaled File System (JFS)**

  Better for lots of small file creates & deletes

  – Buffer caching (default) provides Sequential Read-Ahead, cached writes, etc.
  – Direct I/O (DIO) mount/open option → no caching on reads

- **Enhanced JFS (JFS2)**

  Better for large files/filesystems

  – Buffer caching (default) provides Sequential Read-Ahead, cached writes, etc.
  – Direct I/O (DIO) mount/open option → no caching on reads
  – Concurrent I/O (CIO) mount/open option → DIO, with write serialization disabled
    - **Use for Oracle .dbf, control files and online redo logs only!!!**

- **GPFS**

  Clustered filesystem – the IBM filesystem for RAC

  – Non-cached, non-blocking I/Os (similar to JFS2 CIO) for all Oracle files

> **GPFS and JFS2 with CIO offer similar performance as Raw Devices**

December 5, 2013

AIX Filesystems Mount options (Cont'd)

➢ **Direct IO (DIO)** – introduced in AIX 4.3.

• Data is transfered directly from the disk to the application buffer, bypassing the file buffer cache hence avoiding double caching (filesystem cache + Oracle SGA).
• Emulates a raw-device implementation.

➢ To mount a filesystem in DIO
```
$ mount –o dio /data
```

➢ **Concurrent IO (CIO)** – introduced with JFS2 in AIX 5.2 ML1

• Implicit use of DIO.
• **No Inode locking** : Multiple threads can perform reads and writes on the same file at the same time.
• Performance achieved using CIO is comparable to raw-devices.

➢ To mount a filesystem in CIO:
```
$ mount –o cio /data
```


Bench throughput over run duration – higher tps indicates better performance.

December 5, 2013
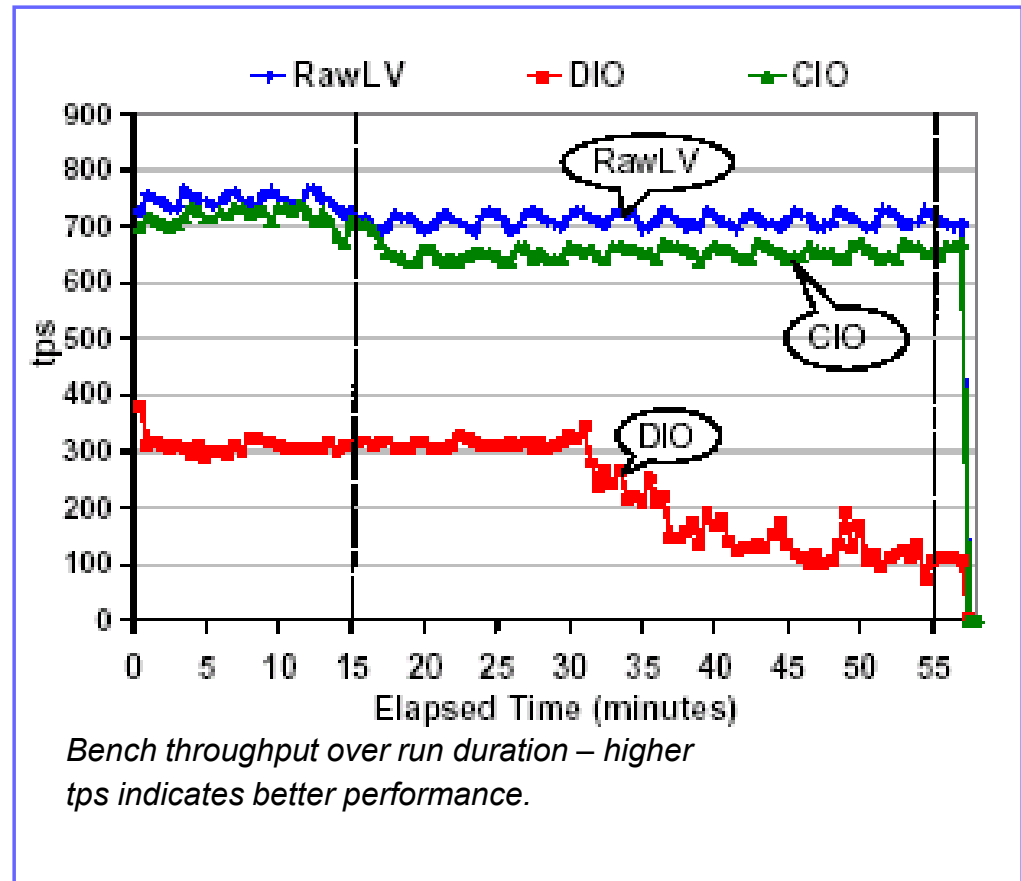
Cached vs. non-Cached (Direct) I/O

File System caching tends to benefit heavily sequential workloads with low write content. **To enable caching for JFS/JFS2:**

- **Use default filesystem mount options**

- **Set Oracle filesystemio_options=ASYNC**

DIO tends to benefit heavily random access workloads and CIO tends to benefit heavy update workloads. **To disable JFS, JFS2 caching, see the following table:**

	Oracle 9i	Oracle 10g/11g
JFS	**Set filesystemio_options=SETALL** **-or-** **Use "dio" mount option**	**Set filesystemio_options=SETALL** **-or-** **Use "dio" mount option**
JFS2	**Use "cio" mount option**	**Set filesystemio_options=SETALL** **-or-** **Use "cio" mount option**

Note: Refer Metalink Note #s 272520.1, 257338.1, 360287.1, 232935.1

December 5, 2013

IBM

CIO/DIO implementation Advices

	with Standard mount options	with optimized mount options
Oracle bin and shared lib.	**mount -o rw** Cached by AIX	**mount -o rw** Cached by AIX
Oracle Datafiles	**mount -o rw** Cached by Oracle Cached by AIX	**mount -o cio** *[1] Cached by Oracle
Oracle Redolog	**mount -o rw** Cached by Oracle Cached by AIX	**mount -o cio (jfs2 + agblksize=512)** Cached by Oracle
Oracle Archivelog	**mount -o rw** Cached by AIX	**mount -o rbrw** Use JFS2 write-behind … but are not kept in AIX Cache.
Oracle Control files	**mount -o rw** Cached by AIX	**mount -o rw** Cached by AIX

Flash Recovery Area mount = rbrw   *(1) : to avoid demoted IO : *jfs2 agblksize = Oracle DB block size / n*

December 5, 2013

CIO Demotion and Filesystem Block Size

Data Base Files (DBF)

- If db_block_size = 2048 → set agblksize=2048

- If db_block_size >= 4096 → set agblksize=4096

Online redolog files & control files

- Set agblksize=512 and use CIO or DIO only for RDBMS 11.1 or less

- Set agblksize=4096 and create redolog with "blocksize" 4096 for RDBMS 11.2

Mount Filesystems with "noatime" option

- AIX/Linux records information about when files were created and last modified as well as last accessed. This may lead to significant I/O performance problems on often accessed files such as the contents of the $ORACLE, /tmp.

I/O Stack Tuning Options (Device Level)

Disk

- **queue_depth** - maximum # of concurrent active I/Os for an hdisk / hdiskpower; additional I/O beyond that limit will be queued. Recommended/supported maximum is storage subsystem dependent.

- **max_transfer** - the maximum allowable I/O transfer size (default is 0x40000 or 256k). Maximum supported value is storage subsystem dependent. All current technology supports 1MB (0x100000) or more I/O size → **set to 0x200000**.

Fiber Channel Disk Adapter *(fcsn)*

- **num_cmd_elems** - maximum number of outstanding I/Os for an adapter.
  set to 1024 or 2048 (**within storage subsystem vendor guidelines)**

- **max_xfer_size** - Increasing value (to at least 0x200000) will also increase DMA size from 16 MB to 256 MB, but this should only be done after IBM support has directed you to do so, as it can lead in specific configurations to system stability issues or AIX not being able to boot.

- **dyntrk** - when set to yes (recommended), allows for immediate re-routing of I/O requests to an alternative path when a device ID (N_PORT_ID) change has been detected; only applies to multi-path configurations.

- **fc_err_recov** - when set to "fast_fail" (recommended), if the driver receives an RSCN notification from the switch, the driver will check to see if the device is still on the fabric and will flush back outstanding I/Os if the device is no longer found.

To validate / change current parameter settings use: "lsattr", "chdev"

AIX IO tuning – LVM Physical buffers (pbuf)

Volume Group (LVM)

AIX

LVM Striping

LV (raw or with jfs2)

hdisk1 — Pbuf

hdisk2 — Pbuf

hdisk3 — Pbuf

hdisk4 — Pbuf

-**Does apply to ASM on top of raw LV !**

-**Absolute value is not important, only increase rate in give time interval**

-**Show lvmo command to get info per VG**
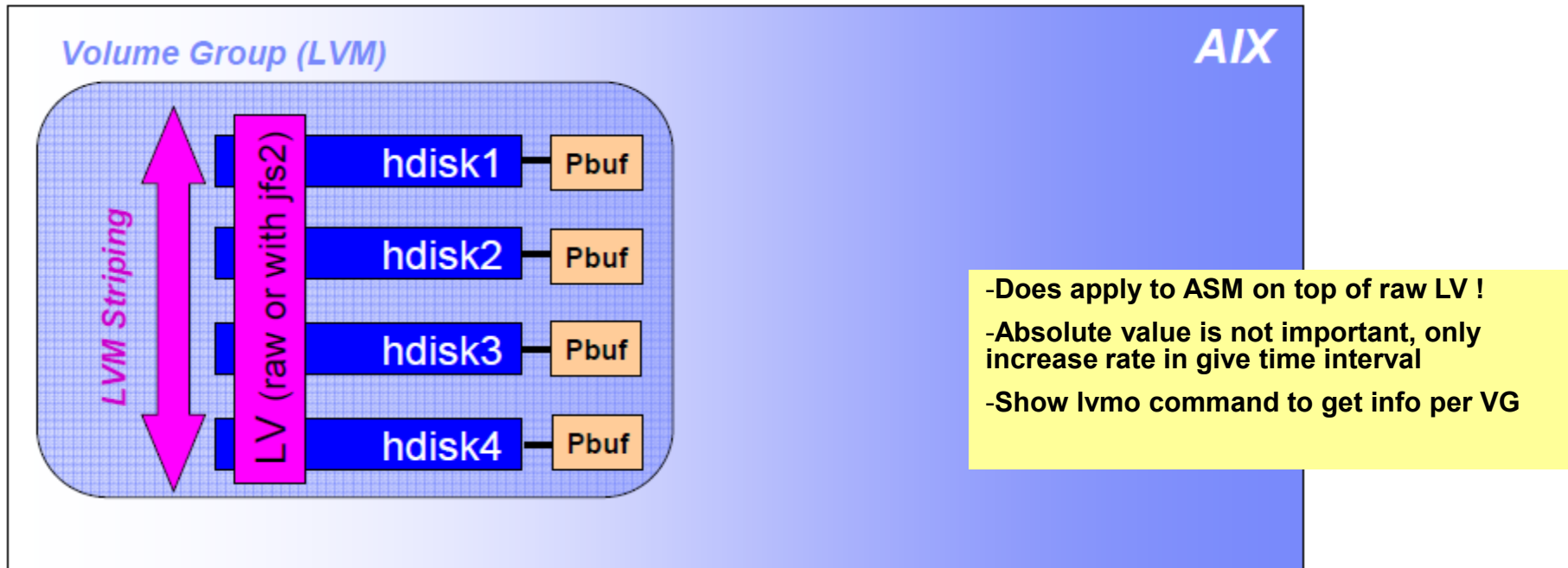
1. Each LVM physical volume as a physical buffer (pbuf)

2. # vmstat –v command help you to detect lack of pbuf.

3. If there is a lack of pbuf, 2 solutions:
   - Add more luns (this will add more pbuf)
   - Or increase pbuf size :
     # lvmo -v <vg_name> -o pv_pbuf_count=XXX

```
0 remote pageouts scheduled
0 pending disk I/Os blocked with no pbuf
0 paging space I/Os blocked with no psbuf
```

Output of "vmstat –v" command

December 5, 2013

AIX IO tuning – HBA tuning (num_cmd_elems)

Volume Group (LVM)

LVM Striping

LV (raw or with jfs2)

hdisk1	Pbuf	queue
hdisk2	Pbuf	queue
hdisk3	Pbuf	queue
hdisk4	Pbuf	queue

Queue_depth

MPIO sddpcm

num_cmd_elems

fcs0 — queue

fcs1 — queue

AIX

External Storage
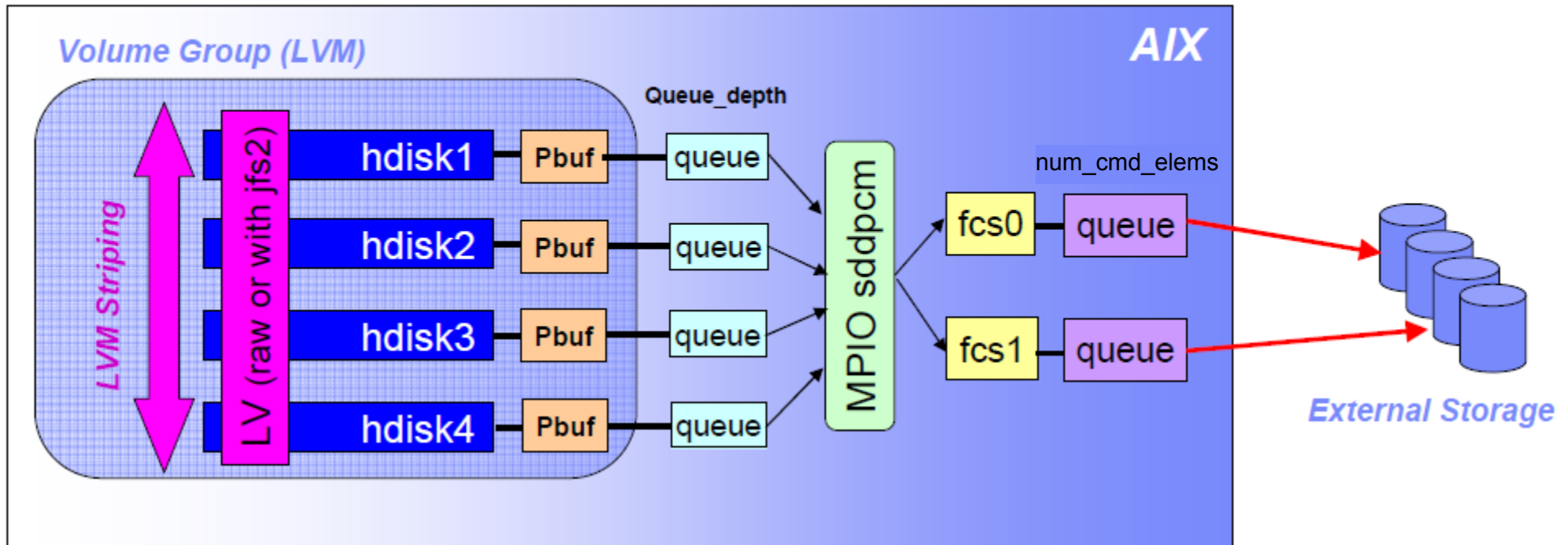
1. Each HBA FC adapter has a queue num_cmd_elems   This queue has the same role for the HBS as the qdepth for the disk.

2. **Rule of thumb:** num_cmd_elems = (sum of qdepth) /**Num HBA**

3. Changing num_cmd_elems : # **chdev -l fcsX -o** num_cmd_elems **=YYY**
   You can also change the max_xfer_size=0x200000
   and lg_term_dma=0x800000 with the same command

```
FC SCSI Adapter Driver Information
  No DMA Resource Count: 0
  No Adapter Elements Count: 0
  No Command Resource Count: 0
```

fcstat fcs0 output

These changes use more memory and must be made with caution, check first with : # fcstat fcsX

Looking for Buffer Structure Shortages

vmstat -v |tail -5   we only need to look at the last 5 lines

- ➤ 0 pending disk I/Os blocked with no pbuf
  if blocked on pbuf, increase
  pv_min_pbuf (ioo restricted) and
  varyoff/varyon VG

- ➤ 0 paging space I/Os blocked with no psbuf
  if blocked on psbuf, stop paging or
  add more paging spaces

- ➤ 2484 filesystem I/Os blocked with no fsbuf
  if blocked on fsbuf (JFS), increase
  numfsbufs (ioo restricted) to 1568

- ➤ 0 client filesystem I/Os blocked with no fsbuf
  if blocked on client fsbuf (NFS) ,
  increase nfso nfs_vX_pdts and
  nfs_vX_vm_bufs values ("X" = 2,3, or 4)

- ➤ 0 external pager filesystem I/Os blocked with no fsbuf
  if blocked on JFS2 fsbuf,
  1) increase j2_dynamicBufferPreallocation (ioo) to 32 (or higher)
  2) If that is not sufficient, increase j2_nBufferPerPagerDevice (ioo restricted)
     to 2048 and unmount/remount JFS2 filesystems

Collect "vmstat –v" output at multiple times and compare statistics –

Look for the values that increase significantly over time

December 5, 2013

IO : Asynchronous IO (AIO)

• Allows multiple requests to be sent without to have to wait until the disk subsystem has completed the physical IO.

• Utilization of asynchronous IO is strongly advised whatever the type of file-system and mount option implemented (JFS, JFS2, CIO, DIO).
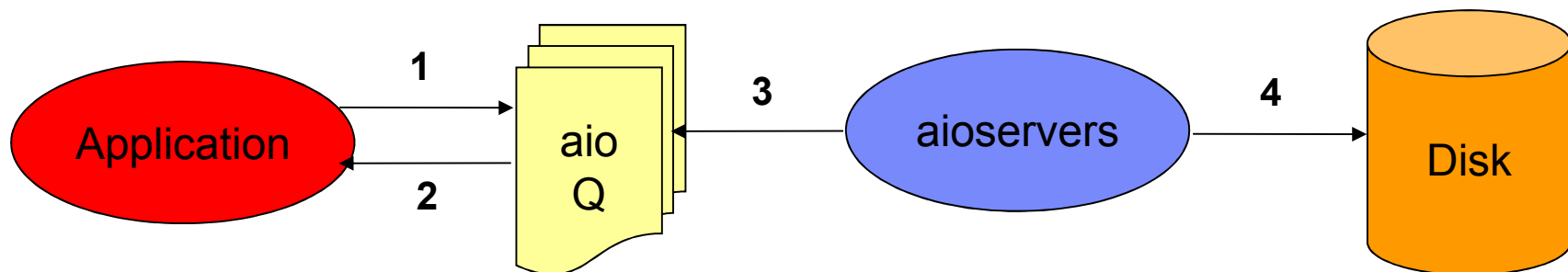
➢Posix vs Legacy

Since AIX5L V5.3, two types of AIO are now available : Legacy and Posix. For the moment, the Oracle code is using the Legacy AIO servers.

December 5, 2013

IO : Asynchronous IO (AIO) fastpath

With fast_path, IO are queued directly from the application into the LVM layer without any "aioservers kproc" operation.

➢ Better performance compare to non-fast_path
➢ No need to tune the min and max aioservers
➢ No ioservers proc. => "ps –k | grep aio | wc –l" is not relevent, use "iostat –A" instead

Application

1

2

AIX Kernel

3

Disk
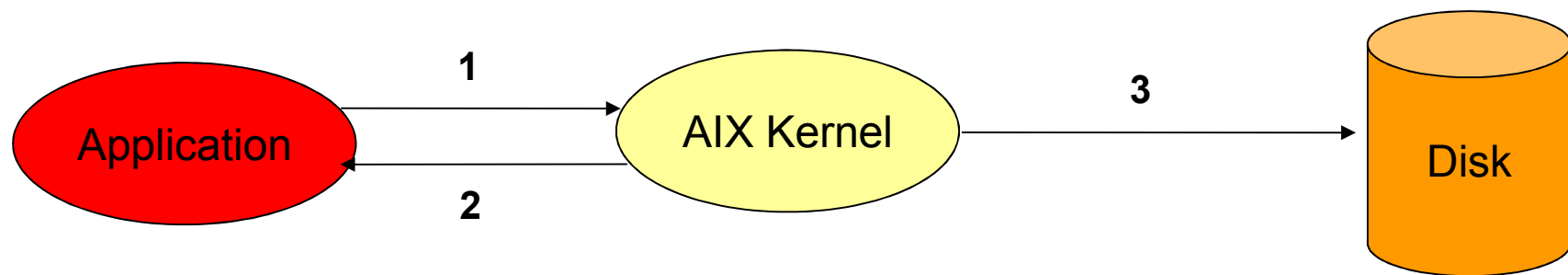
• **Raw Devices / ASM :**
➢ check AIO configuration with : lsattr –El aio0
enable asynchronous IO **fast_path. :**
AIX 5L : chdev -a fastpath=enable -l aio0 *(default since AIX 5.3)*
AIX 6.1/7.1 : ioo –p –o aio_fastpath=1   *(default setting)*

• **FS with CIO/DIO and AIX 5.3 TL5+ :**
➢ Activate **fsfast_path** (comparable to fast_path but for FS + CIO/DIO)
AIX 5L : adding the following line in /etc/inittab: **aioo:2:once:aioo –o fsfast_path=1**
AIX 6.1/7.1 : ioo –p –o aio_fsfastpath=1 *(default setting)*

December 5, 2013

Asynchronous I/O for <u>filesystem</u> environments…

Monitor Oracle usage:

- **Watch alert log and *.trc files in BDUMP directory for warning message:**

     "Warning "lio_listio returned EAGAIN"

➤ If warning messages found, increase maxreqs and/or maxservers

Monitor from AIX:

- "pstat –a | grep aios"

- Use "-A" option for NMON

- iostat –Aq (new in AIX 5.3)

December 5, 2013

GPFS I/O Related Tunables

- **Refer Metalink note 302806.1**

Async I/O:

- Oracle parameter filesystemio_options is ignored
- Set Oracle parameter disk_asynch_io=TRUE
- Prefetchthreads= exactly what the name says
  - Usually set prefetchthreads=64 (the default)
- Worker1threads = GPFS asynch I/O
  - Set worker1threads=550-prefetchthreads
- Set aio maxservers=(worker1threads/#cpus) + 10

Other settings:

- GPFS block size is configurable; most will use 512KB-1MB
- **Pagepool –** GPFS fs buffer cache, not used for RAC but may be for binaries.   Default=64M

mmchconfig pagepool=100M

AIX rendev command with ASM

rendev command is used for renaming devices which are listed in ODM

- **Syntax / Description**
  - rendev –l <original name> -n <new name>
  - The device entry under /dev will be renamed corresponding to <new name>
  - Certain devices such as /dev/console, /dev/mem, /dev/null, and others that are identified only with /dev special files cannot be renamed
  - Command will fail for any device that does not have both a *Configure* and an *Unconfigure* method
  - Any name that is 15 characters or less and not already used in the system can be used

- **If used to rename hdisk devices for ASM use, it is recommended that you keep the "hdisk" prefix, as this will allow the default ASM discovery string to match the renamed hdisks. Corresponding rhdisk is renamed as well.**

  Example:

  ```
  # rendev –l hdisk10 –n hdiskASM10
  # ls /dev/*ASM*
  /dev/hdiskASM10
  /dev/rhdiskASM10
  ```

AIX lkdev command with ASM

- **The lkdev command locks the specified device. Any attempt to modify device attributes by using the chdev or chpath command is denied. In addition, an attempt to delete the specified device or one of its paths from the ODM by using either the rmdev or rmpath command is denied.**

- **Syntax:**

  lkdev [-l <Name> -a | -d [-c <Text>]]

<Name>	Name of device to be changed (required)
-a	Locks the specified device.
-d	Unlocks the specified device.
-c <Text>	Specifies a text of up to 64 printable characters with no embedded spaces.

- **Examples:**

  – To enable the lock for the *hdiskASM10* disk device and create a text label, enter the following command:

    ```
    # lkdev -l hdiskASM10 -a -c ASMdisk
    ```

  – To remove the lock for the *hdisk1* disk device and remove the text label, enter the following command:

    ```
    # lkdev -l hdiskASM10 -d
    ```

    Note: The text label of a locked device can not be changed! Instead, the device needs to be first unlocked and then locked again with the new text label specified.

September 2012

December 5, 2013

ASM configurations

AIX parameters

– Async I/O needs to be enabled, but default values may be used

ASM instance parameters

– ASM_POWER_LIMIT=1

Makes ASM rebalancing a low-priority operation. May be changed dynamically.   It is common to set this value to 0, then increase to a higher value during maintenance windows

– PROCESSES=25+ 15n, where n=# of instances using ASM

DB instance parameters

– disk_asynch_io=TRUE

– filesystemio_options=ASYNCH

– Increase Processes by 16

– Increase Large_Pool by 600k

– Increase Shared_Pool by [(1M per 100GB of usable space) + 2M]

I/O Pacing

- **I/O Pacing parameters can be used to prevent large I/O streams from monopolizing CPUs**

  - **System backups (mksysb)**

  - **DB backups (RMAN, Netbackup)**

  - **Software patch updates**

- **When Oracle ClusterWare is used, use AIX 6.1/7.1 Defaults:**

  - **chgsys -l sys0 -a maxpout=8193 –a minpout=4096 (AIX defaults)**

  - **nfso –o nfs_iopace_pages=1024 (AIX defaults)**

  - **On the Oracle clusterware set : crsctl set css diagwait 13 –force**

    - This will delay the OPROCD reboot time to 10secs from 0.5secs during node eviction/reboot, just enough to write the log/trace files for future diagnosis. Metalink note# 559365.1

December 5, 2013

Agenda

- **AIX Configuration Best Practices for Oracle**

  – Memory

  – CPU

  – I/O

  – **Network**

  – Miscellaneous

December 5, 2013

Grid Interconnect Redundancy – Recommended

- **Etherchannel (802.3ad)**

  - Up to 8 primary adapters per etherchannel

  - All adapters in the etherchannel should be the same, and should be configured identically (e.g., gigabit full duplex)

  - All adapters constituting an etherchannel must be connected to the same switch

  - Switch support for etherchannel (may be called "aggregation" or "trunking") required

- **Network Interface Backup**

  - One per etherchannel

  - No switch requirements

  - Provides backup capability only, no increased throughput like etherchannel

- **Shared Ethernet Adapter (SEA) Failover**

  - VIOS based failover and load balancing

- **New in 11.2.0.2 – Oracle HAIP**

  - Up to 4 primary adapters can be aggregated for redundancy and bandwidth

  - Requires multicast communication on 230.0.1.0 network, although 224.0.0.251 will work with patch 9974223

  - 11.2.0.3 HAIP will work with broadcast communication

Network Options (no) Parameters

- **use_isno = 1 means any parameters set at the interface level override parameters set with 'no'**

  – DEFAULT (restricted) in AIX 6.1

- **If use_isno = 0, any parameters set with 'no' override interface-specific parameters**

- **If use_isno = 1, set parameters for each interface using 'ifconfig' or 'chdev'**

- **Refer to the following URL for a chart on appropriate interface-specific parameters:**

  – http://publib.boulder.ibm.com/infocenter/systems/topic/com.ibm.aix.prftungd/doc/prftungd/prftungd.pdf

- **Generally appropriate parameters for 1 or 10 Gigabit Ethernet Oracle public network interfaces:**

  – tcp_sendspace = 262144

  – tcp_recvspace = 262144

  – rfc1323 = 1

- **Examples:**

  – # no -p -o tcp_sendspace=262144

  – # no -p -o tcp_recvspace=262144

  – # no -p -o rfc1323=1

December 5, 2013

TCP/IP Ephemeral Ports

- **Oracle 11gR2 checks the current TCP/IP ephemeral port range at binary install time**

  – If the current tcp and udp ephemeral port ranges are not set to 9000 – 65500, the Oracle installer will generate warning messages

- **If the default range (32768 – 65535) is sufficient to support the anticipated server workload, the warnings may be ignored**

- **If the workload will require a high number of ephemeral ports, such as high node counts or heavy use of Parallel Query – or to avoid the install time warning messages, the ephemeral port ranges may be re-configured**

- **Examples:**

  – # no -p -o tcp_ephemeral_low=9000

  – # no -p -o tcp_ephemeral_high=65500

  – # no -p -o udp_ephemeral_low=9000

  – # no -p -o udp_ephemeral_high=65500

December 5, 2013

Additional Network (no) Parameters for RAC:

- Set udp_sendspace = db_block_size * db_file_multiblock_read_count

  (not less than 65536)

- Set udp_recvspace = 10 * udp_sendspace

  – Must be < sb_max

- Increase if buffer overflows occur

- Ipqmaxlen=512 for GPFS environments

- Use Jumbo Frames if supported at the switch layer

Examples:

- no -a |grep udp_sendspace

- no –o -p udp_sendspace=65536

- netstat -s |grep "socket buffer overflows"

December 5, 2013

Agenda

- **AIX Configuration Best Practices for Oracle**

  – Memory

  – I/O

  – Network

  – **Miscellaneous**

Miscellaneous parameters

- **User Limits (smit chuser) /etc/security/limits**
  - Soft **FILE** size = -1 (Unlimited)
  - Soft **CPU** time = -1 (Unlimited)
  - Soft **DATA** segment = -1 (Unlimited)
  - Soft **STACK** size -1 (Unlimited)

- **Sys0 maxuproc attribute**
  - Should be >= 16384
  - For workloads with a large number of concurrent connections an/or parallel servers, should be at least 128 plus the sum of PROCESSES and PARALLEL_MAXSERVERS for all instances in the LPAR

- **Environment variables:**
  - AIXTHREAD_SCOPE=S
  - LDR_CNTRL=DATAPSIZE=64K@TEXTPSIZE=64K@STACKPSIZE=64K@*SHMPSIZE=64K*

December 5, 2013

Show me the I/O Speed

December 5, 2013

Sample Configuration

RamSan 70

P770 32core 3.55ghz 256gb memory

AIX 7.1 / Oracle 11.2.0.3

PCIe

RamSan 630 / 10TB

Read IOPS 1 M IOPS

Read Latency: 250 µs

Write IOPS: 1 M IOPS

Write Latency: 80 µs

Fiber Channel/ InfiniBand

Data, index

2-4gb Fiber Channel

Fiber Channel/ InfiniBand

Existing SAN storage

Data

Indexes

RamSan 820 / 10-20 TB

Read IOPS 450 K IOPS

Read Latency: 110 µs

Write IOPS: 400 K IOPS

Write Latency: 25 µs

Redo, Undo, temp, System, Sysaux

➢ **Where to find the documentation/ reference/wiki**

December 5, 2013

IBM TechDocs - Technical Sales Library

http://www.ibm.com/support/techdocs

- **Oracle Architecture and Tuning on AIX v2.20**

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100883

- **Configuring IBM TotalStorage for Oracle OLTP Applications**

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100319

- **Diagnosing Oracle® Database Performance on AIX® Using IBM® NMON and Oracle Statspack Reports**

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101720

- **Breaking the Oracle I/O Performance Bottleneck**

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS3885

- **Oracle Technology Essential White Papers**   Regularly Updated!

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101559

There are many more Oracle related white papers – especially covering Oracle RAC with IBM servers and IBM storage.

December 5, 2013

Oracle Documentation

☞ **My Oracle Support:** http://support.oracle.com   `Regularly Updated!`

   – **282036.1 - Minimum Software Versions and Patches Required to Support Oracle Products on IBM Power Systems**

   ▪ **756671.1 - Oracle Recommended Patches -- Oracle Database**

▪ **Oracle Reference Manuals:**

   http://otn.oracle.com/documentation/index.html

▪ **Oracle Certification Info (on MOS as well):**

   http://otn.oracle.com/support/metalink/index.html

December 5, 2013

AIX

- AIX 5.3 Product Documentation.
  - http://publib.boulder.ibm.com/infocenter/pseries/v5r3/index.jsp?topic=/com.ibm.pseries.doc/hardware.htm
- AIX 6.1 Product Documentation
  - http://publib.boulder.ibm.com/infocenter/aix/v6r1/index.jsp
- AIX 7.1 Product Documentation
  - http://publib.boulder.ibm.com/infocenter/aix/v7r1/index.jsp
  - http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=sg247910 (IBM AIX Version 7.1 Differences Guide)

- **IBM Wikis**
  - https://www.ibm.com/developerworks/wikis/dashboard.action
- AIX Wiki
  - http://www.ibm.com/developerworks/wikis/display/WikiPtype/Home
- AIX Performance Tools (nmon, nmon analyser/consolidator, etc)
  - http://www.ibm.com/developerworks/wikis/display/WikiPtype/nmon
- AIX DeveloperWorks
  - http://www.ibm.com/developerworks/aix
- AIX multiple page supprt
  - http://www-03.ibm.com/systems/resources/systems_p_os_aix_whitepapers_multiple_page.pdf
- Tuning IBM AIX 5L V5.3 and AIX 6.1 for Oracle Database on POWER systems
  - http://www-304.ibm.com/partnerworld/wps/servlet/ContentHandler/whitepaper/aix/oracle/performance_analysis
- PowerVM Wiki
  - https://www.ibm.com/developerworks/wikis/display/virtualization/Home

AIX / POWER

- **AIXpert Blog on Local, Near and Far Memory**

  https://www.ibm.com/developerworks/mydeveloperworks/blogs/aixpert/entry/local_near_far_memory_part_1_large_power7_boxes_more_local_memory26?lang=en

- **Oracle Database and 1 TB Segment Aliasing (TD105761)**

  http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105761

- **IBM EnergyScale for POWER7 Processor-Based Systems**

  ftp://public.dhe.ibm.com/common/ssi/ecm/en/pow03039usen/POW03039USEN.PDF

- **Active Memory Expansion: Overview and Usage Guide**

  ftp://ftp.software.ibm.com/common/ssi/sa/wh/n/pow03037usen/POW03037USEN.PDF

- **IBM PowerVM Virtualization Active Memory Sharing**

  http://www.redbooks.ibm.com/abstracts/redp4470.html?Open

- **IBM System p Advanced POWER Virtualization (PowerVM) Best Practices**

  http://www.redbooks.ibm.com/abstracts/redp4194.html?Open

- **Power Systems Enterprise Servers with PowerVM Virtualization and RAS**

  http://www.redbooks.ibm.com/abstracts/sg247965.html?Open

December 5, 2013

IBM

Thank You

December 5, 2013

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/30, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, z/VM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries
LINUX is a registered trademark of Linux Torvalds
UNIX is a registered trademark of The Open Group in the United States and other countries.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
Intel is a registered trademark of Intel Corporation
* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput 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 improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

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

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

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

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

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

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.