

AIX PERFORMANCE TUNING

This presentation at:

<http://www.circle4.com/papers/common-performance.pdf>



Jaqui Lynch
lynchj@forsythe.com



AGENDA

- CPU
- Memory tuning
- Network
- Starter Set of Tunables
- Volume groups and filesystems
- AIO and CIO for Oracle

- Backup Slides
 - Performance Tools



CPU

3



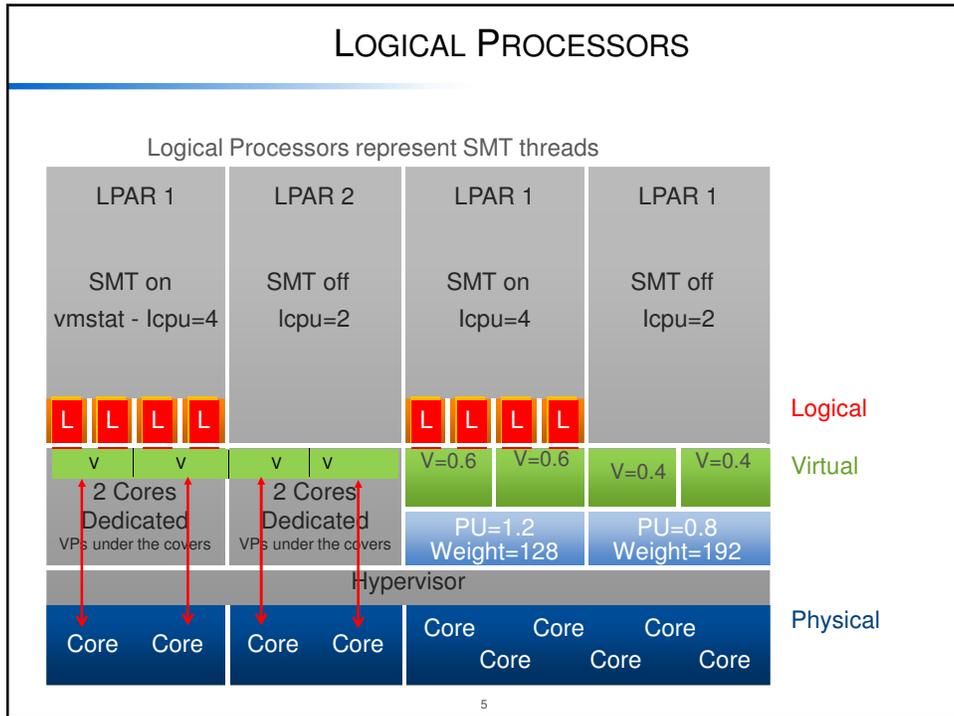
MONITORING CPU

User, system, wait and idle are fine for dedicated LPARs
They are not fine for SPLPAR or dedicated donating LPARs
You need to measure and charge back based on used CPU cycles
Moral of the story – use Physc (Physical consumed)
lparstat

- Use with no flags to view partition configuration and processor usage

4





UNDERSTAND SMT

SMT

- Threads dispatch via a Virtual Processor (VP)
- SMT1: Largest unit of execution work
- SMT2: Smaller unit of work, but provides greater amount of execution work per cycle
- SMT4: Smallest unit of work, but provides the maximum amount of execution work per cycle
- On POWER7, a single thread cannot exceed 65% utilization
- On POWER6 or POWER5, a single thread can consume 100%
- Understand thread dispatch order
- VPs are unfolded when threshold is reached
 - P5 and P6 primary and secondary threads are loaded to 80% before another VP unfolded
 - In P7 primary threads are loaded to 50%, unfolds VPs, then secondary threads used. When they are loaded to 50% tertiary threads are dispatched

SMT Thread

Primary 0

Secondary 1

Tertiary 2, 3

Diagram courtesy of IBM

FORSYTHE

6

UNDERSTAND YOUR WORKLOAD

Are you about speed?

- Speed is distance over time or performance
- Affected by clock speed, memory and I/O bandwidth, etc
- Basically how much can I push through one core
- Higher frequency cores
- May run better with SMT2 or SMT or dedicated cores

Or throughput?

- Volume over time or capacity
- How many concurrent things can I push through
- Affected by pipelining and SMT

Architect accordingly

Check for gating factors that could impact use of SMT

- i.e. is there one thread that controls all work?

7



APPLICATIONS AND SPLPARS

Applications do not need to be aware of Micro-Partitioning

Not all applications benefit from SPLPARS

Applications that may not benefit from Micro-Partitioning:

- Applications with a strong response time requirements for transactions may find Micro-Partitioning detrimental:
 - Because virtual processors can be dispatched at various times during a timeslice
 - May result in longer response time with too many virtual processors:
 - Each virtual processor with a small entitled capacity is in effect a slower CPU
 - Compensate with more entitled capacity (2-5% PUs over plan)
- Applications with polling behavior
- CPU intensive application examples: DSS, HPC, SAS

Applications that are good candidates for Micro-Partitioning:

- Ones with low average CPU utilization, with high peaks:
 - Examples: OLTP, web applications, mail server, directory servers

In general Oracle databases are fine in the shared processor pool

For licensing reasons you may want to use a separate pool for databases

8



DISPATCHING IN SHARED POOL

VP gets dispatched to a core

- First time this becomes the home node
- All 4 threads for the VP go with the VP

VP runs to the end of its entitlement

- If it has more work to do and noone else wants the core it gets more
- If it has more work to do but other VPs want the core then it gets context switched and put on the home node runQ
- If it can't get serviced in a timely manner it goes to the global runQ and ends up running somewhere else but its data may still be in the memory on the home node core

9



MORE ON DISPATCHING

Nigel Griffiths Power7 Affinity – Session 19 and 20 - <http://tinyurl.com/newUK-PowerVM-VUG>



10



ENTITLEMENT AND VPS

Utilization calculation for CPU is different between POWER5, 6 and POWER7
VPs are also unfolded sooner (at lower utilization levels than on P6 and P5)

This means that in POWER7 you need to pay more attention to VPs

- You may see more cores activated at lower utilization levels
- But you will see higher idle
- If only primary SMT threads in use then you have excess VPs

Try to avoid this issue by:

- Reducing VP counts
- Use realistic entitlement to VP ratios
 - 10x or 20x is not a good idea
 - Try setting entitlement to .6 or .7 of VPs
- Ensure workloads never run consistently above 100% entitlement
- Too little entitlement means too many VPs will be contending for the cores
- **Performance may (in most cases, will) degrade when the number of Virtual Processors in an LPAR exceeds the number of physical processors**

11



SCALED THROUGHPUT

P7 and P7+ with AIX v6.1 TL08 and AIX v7.1 TL02

Dispatches more SMT threads to a VP core before unfolding additional VPs

Tries to make it behave a bit more like P6

Raw provides the highest per-thread throughput
and best response times at the expense of activating more physical core

Scaled provides the highest core throughput at the expense of per-thread response times and throughput.

It also provides the highest system-wide throughput per VP because tertiary thread capacity is "not left on the table."

schedo -p -o vpm_throughput_mode=

- 0 Legacy Raw mode (default)
 - 1 "Enhanced Raw" mode with a higher threshold than legacy
 - 2 Scaled mode, use primary and secondary SMT threads
 - 4 Scaled mode, use all four SMT threads
- Dynamic Tunable

SMT unfriendly workloads could see an enormous per thread performance degradation

12



USEFUL PROCESSOR COMMANDS

```
lsdev -Cc processor
lsattr -EL proc0
bindprocessor -q
sar -P ALL
topas, nmon
lparstat
vmstat (use -l or -v)
iostat
mpstat -s
lssrad -av
```

13



USING SAR -P ALL (POWER7 & SMT4)

AIX bpicnim 1 7 00F6934B4C00 10/05/11 (1 core and 2 VPs)

System configuration: lcpu=8 ent=1.00 mode=Uncapped

```
19:40:49 cpu %usr %sys %wio %idle physc %entc
19:40:50  0  7  88  0  5  0.01  1.4
          1  0  0  0 100  0.00  0.3
          2  0  1  0  99  0.00  0.3
          3  0  0  0 100  0.00  0.3
          7  0 59  0  41  0.00  0.0
          U  -  -  0  98  0.98 97.5
          -  0  1  0  99  0.02  2.5
```

In the above cpu4-6 are missing as they are 0 so sar did not print them to save space

mpstat -s 1 1

System configuration: lcpu=8 ent=1.0 mode=Uncapped

```

Proc0
2.26%
cpu0  cpu1  cpu2  cpu3
1.33% 0.31% 0.31% 0.31%

Proc4
0.01%
cpu4  cpu5  cpu6  cpu7
0.00% 0.00% 0.00% 0.01%
```

14



USING SAR -P ALL - AIX 5.3 OR SMT2

SMT2 Example

```
sar -P ALL 1 1
AIX sys01a 3 5 00CDAF6F4C00 ent=0.80
System Configuration: lcpu=4 ent=0.80
12:18:01  cpu    %usr  %sys  %wio  %idle  physc  %entc
12:18:01  0      0     7     0     93     0.03   3.3
          1     100   0     0     0     0.37   46.8
          2     100   0     0     0     0.38   46.9
          3      0     1     0     99     0.02   3.1
          -     94    0     0     6     0.80  100
```

physc total matches ent above so 100%

System is clearly busy - now map this to the mpstat command

```
mpstat -s 1 1
System configuration: lcpu=4 ent=0.80
          Proc0                      Proc1
          39.99%                      39.76%
          cpu0          cpu1          cpu2          cpu3
          2.55%         37.45%         37.57%         2.19%
```

Oracle tends to really like SMT and to take advantage of it

15



MEMORY

16



MEMORY TYPES

Persistent

- Backed by filesystems

Working storage

- Dynamic
- Includes executables and their work areas
- Backed by page space

Prefer to steal from persistent as it is cheap
 minperm, maxperm, maxclient, lru_file_repage and
 page_steal_method all impact these decisions

17



CORRECTING PAGING

11173706 paging space I/Os blocked with no psbuf

Isps output on above system that was paging before changes were made to tunables

Isps -a

Page Space	Physical Volume	Volume Group	Size	%Used	Active	Auto	Type
paging01	hdisk3	pagingvg	16384MB	25	yes	yes	lv
paging00	hdisk2	pagingvg	16384MB	25	yes	yes	lv
hd6	hdisk0	rootvg	16384MB	25	yes	yes	lv

What you want to see

Isps -a

Page Space	Physical Volume	Volume Group	Size	%Used	Active	Auto	Type
paging01	hdisk3	pagingvg	16384MB	1	yes	yes	lv
paging00	hdisk2	pagingvg	16384MB	1	yes	yes	lv
hd6	hdisk0	rootvg	16384MB	1	yes	yes	lv

Isps -s

Total Paging Space	Percent Used	Can also use vmstat -l and vmstat -s
49152MB	1%	

Should be balanced – NOTE VIO Server comes with 2 different sized page datasets on one hdisk (at least until FP24)

18



DEFAULT PAGE SPACE CALCULATION

AIX Client default

- hd6 must be $\geq 64\text{MB}$, others must be $\geq 16\text{MB}$
- Page space can use no more than 20% disk
- If real $< 256\text{MB}$ then page space = $2 \times \text{real}$
- If real $\geq 256\text{MB}$ then page space = 256MB

VIO Server

- 1 x 512MB and 1 x 1024MB page space both on the same disk
- Supposedly fixed if installing FP24 but not if upgrade

On my VIO:

```
# lspv -a
```

Page Space	Physical Volume	Volume Group	Size	%Used	Active	Auto	Type	Chksum
hd6	hdisk0	rootvg	4096MB	1	yes	yes	lv	0

19



PAGE SPACE BEST PRACTICE

More than one page volume

All the same size including hd6

Page spaces must be on different disks to each other

Do not put on hot disks

Mirror all page spaces that are on internal or non-raided disk

20



MEMORY WITH LRU_FILE_REPAGE=0

minperm=3

- Always try to steal from filesystems if filesystems are using more than 3% of memory

maxperm=90

- Soft cap on the amount of memory that filesystems or network can use
- Superset so includes things covered in maxclient as well

maxclient=90

- Hard cap on amount of memory that JFS2 or NFS can use – SUBSET of maxperm

21



PAGE_STEAL_METHOD

Default in 5.3 is 0, in 6 and 7 it is 1

What does 1 mean?

lru_file_repage=0 tells LRUD to try and steal from filesystems

Memory split across mempools

LRUD manages a mempool and scans to free pages

0 – scan all pages

1 – scan only filesystem pages

22



PAGE_STEAL_METHOD EXAMPLE

500GB memory
 50% used by file systems (250GB)
 50% used by working storage (250GB)
 mempools = 5
 So we have at least 5 LRUDs each controlling about 100GB memory
 Set to 0

- Scans all 100GB of memory in each pool

Set to 1

- Scans only the 50GB in each pool used by filesystems

Reduces cpu used by scanning
 When combined with CIO this can make a significant difference

23



LSSRAD -AV

Large LPAR on a 770

REF1	SRAD	MEM	CPU
0	0	171529.00	0-15 20-23 28-31 36-39 44-47 52-55 60-
63	1	114771.50	16-19 24-27 32-35 40-43 48-51 56-59

REF1 indicates where
 REF1=0 SRAD=0 is local
 REF1=0 SRAD=1 is near
 Other REF values are far
 This is relative to the process home

Smaller LPAR

REF1	SRAD	MEM	CPU
0	0	88859.50	0-7
	2	36354.00	
1	1	42330.00	8-11
	3	20418.00	

24



STARTER SET OF TUNABLES 1

For AIX v5.3

No need to set memory_affinity=0 after 5.3 tl05

MEMORY

vmo -p -o minperm%=3

vmo -p -o maxperm%=90

vmo -p -o maxclient%=90

vmo -p -o minfree=960

We will calculate these

vmo -p -o maxfree=1088

We will calculate these

vmo -p -o lru_file_repage=0

vmo -p -o lru_poll_interval=10

vmo -p -o page_steal_method=1

For AIX v6 or v7

Memory defaults are already correctly except minfree and maxfree

If you upgrade from a previous version of AIX using migration then you need to check the settings after

25



STARTER SET OF TUNABLES 2

The parameters below should be reviewed and changed (see vmstat -v and lvmo -a later)

PBUFS

Use the new way (coming up)

JFS2

ioo -p -o j2_maxPageReadAhead=128

(default above may need to be changed for sequential)

j2_dynamicBufferPreallocation=16

Default that may need tuning

Replaces tuning j2_nBufferPerPagerDevice

Network changes in later slide

26



ROUGH ANATOMY OF AN I/O

LVM requests a PBUF

- Pinned memory buffer to hold I/O request in LVM layer

Then placed into an FSBUF

- 3 types
- These are also pinned
- Filesystem
- Client
- External Pager

JFS
NFS and VxFS
JFS2

If paging also need PSBUFs (also pinned)

- Used for I/O requests to and from page space

Then queue I/O to hdisk (queue_depth)

Then queue it to adapter (num_cmd_elems)

Adapter queues it to the disk subsystem

27



LVMO -A OUTPUT

2725270 pending disk I/Os blocked with no pbuf

vgname = rootvg

pv_pbuf_count = 512

total_vg_pbufs = 1024

max_vg_pbuf_count = 16384

pervg_blocked_io_count = 0

this is rootvg

pv_min_pbuf = 512

Max_vg_pbuf_count = 0

global_blocked_io_count = 2725270

this is the others

Use lvmo -v xxxvg -a

For other VGs we see the following in pervg_blocked_io_count

	blocked	total_vg_bufs
nimvg	29	512
sasvg	2719199	1024
backupvg	6042	4608

lvmo -v sasvg -o pv_pbuf_count=2048

28



VMSTAT -V OUTPUT

3.0 minperm percentage
 90.0 maxperm percentage
 45.1 numperm percentage
 45.1 numclient percentage
 90.0 maxclient percentage

1468217 pending disk I/Os blocked with no pbuf	pbufs
11173706 paging space I/Os blocked with no psbuf	pagespace
2048 file system I/Os blocked with no fsbuf	JFS
238 client file system I/Os blocked with no fsbuf	NFS/VxFS
39943187 external pager file system I/Os blocked with no fsbuf	JFS2

numclient=numperm so most likely the I/O being done is JFS2 or NFS or VxFS
 Based on the blocked I/Os it is clearly a system using JFS2
 It is also having paging problems
 pbufs also need reviewing

29



MEMORY POOLS AND FRE COLUMN

fre column in vmstat is a count of all the free pages across all the memory pools

When you look at fre you need to divide by memory pools

Then compare it to maxfree and minfree

This will help you determine if you are happy, page stealing or thrashing

You can see high values in fre but still be paging

In below if maxfree=2000 and we have 10 memory pools then we only have 990 pages free in each pool on average. With minfree=960 we are page stealing and close to thrashing.

kthr		memory				page				faults				cpu			
r	b	p	avm	fre	fi	fo	pi	po	fr	sr	in	sy	cs	us	sy	id	wa
70	309	0	8552080	9902	75497	9615	9	3	84455	239632	18455	280135	91317	42	37	0	20

9902/10 = 990.2

30



CALCULATING MINFREE AND MAXFREE

```
vmstat -v | grep memory
      3 memory pools
```

```
vmo -a | grep free
      maxfree = 1088
      minfree = 960
```

Calculation is:

$$\text{minfree} = (\max(960, (120 * \text{lcpus}) / \text{memory pools}))$$

$$\text{maxfree} = \text{minfree} + (\text{Max}(\text{maxpgahead}, \text{j2_maxPageReadahead}) * \text{lcpus}) / \text{memory pools}$$

So if I have the following:

Memory pools = 3 (from vmo -a or kdb)

J2_maxPageReadahead = 128

CPUS = 6 and SMT on so lcpu = 12

So minfree = $(\max(960, (120 * 12) / 3)) = 1440 / 3 = 480$ or 960 whichever is larger

And maxfree = $\text{minfree} + (128 * 12) / 3 = 960 + 512 = 1472$

I would probably bump this to 1536 rather than using 1472 (nice power of 2)

If you over allocate these values it is possible that you will see high values in the "fre" column of a vmstat and yet you will be paging.

31



NETWORK

32



STARTER SET OF TUNABLES 3

Typically we set the following for both versions:

NETWORK

```
no -p -o rfc1323=1
no -p -o tcp_sendspace=262144
no -p -o tcp_recvspace=262144
no -p -o udp_sendspace=65536
no -p -o udp_recvspace=655360
```

Also check the actual NIC interfaces and make sure they are set to at least these values

Check sb_max is at least 1040000 – increase as needed

33



IFCONFIG

ifconfig -a output

```
en0:
flags=1e080863,480<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,
T,GROUPRT,64BIT,CHECKSUM_OFFLOAD(ACTIVE),CHAIN>
    inet 10.2.0.37 netmask 0xffffe00 broadcast 10.2.1.255
    tcp_sendspace 65536 tcp_recvspace 65536 rfc1323 0

lo0:
flags=e08084b<UP,BROADCAST,LOOPBACK,RUNNING,SIMPLEX,MULTICAST,GROU
PRT,64BIT>
    inet 127.0.0.1 netmask 0xff000000 broadcast 127.255.255.255
    inet6 ::1/0
    tcp_sendspace 131072 tcp_recvspace 131072 rfc1323 1
```

These override no, so they will need to be set at the adapter.
Additionally you will want to ensure you set the adapter to the correct setting if it runs at less than GB, rather than allowing auto-negotiate

Stop inetd and use chdev to reset adapter (i.e. en0)
Or use chdev with the -P and the changes will come in at the next reboot

34



MY VIO SERVER SEA

```

# ifconfig -a
en6:
flags=1e080863,580<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,M
ULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD(ACTIVE),CHAIN>

    inet 192.168.2.5 netmask 0xfffff00 broadcast 192.168.2.255
    tcp_sendspace 262144 tcp_recvspace 262144 rfc1323 1

lo0:
flags=e08084b,1c0<UP,BROADCAST,LOOPBACK,RUNNING,SIMPLEX,MUL
TICAST,GROUPRT,64BIT,LARGESEND,CHAIN>
    inet 127.0.0.1 netmask 0xff000000 broadcast 127.255.255.255
    inet6 ::1%1/0
    tcp_sendspace 131072 tcp_recvspace 131072 rfc1323 1
    
```



NETWORK

Interface	Speed	MTU	tcp_sendspace	tcp_recvspace	rfc1323
lo0	N/A	16896	131072	131072	1
Ethernet	10/100 mb				
Ethernet	1000 (Gb)	1500	131072	165536	1
Ethernet	1000 (Gb)	9000	262144	131072	1
Ethernet	1000 (Gb)	1500	262144	262144	1
Ethernet	1000 (Gb)	9000	262144	262144	1
Virtual Ethernet	N/A	any	262144	262144	1
InfiniBand	N/A	2044	131072	131072	1

Above taken from Page 247 SC23-4905-04 November 2007 edition
 Check up to date information at:
<http://publib.boulder.ibm.com/infocenter/pseries/v5r3/topic/com.ibm.aix.prfungd/doc/prftungd/prftungd.pdf>

AIX v6.1
http://publib.boulder.ibm.com/infocenter/aix/v6r1/topic/com.ibm.aix.prfungd/doc/prftungd/prftungd_pdf.pdf



NETWORK COMMANDS

- entstat -d or netstat -v (also -m and -l)
- netpmon
- iptrace (traces) and ipreport (formats trace)
- Tcpdump
- traceroute
- chdev, lsattr
- no
- ifconfig
- ping and netperf
- ftp
 - Can use ftp to measure network throughput
 - ftp to target
 - ftp> put "| dd if=/dev/zero bs=32k count=100" /dev/null
 - Compare to bandwidth (For 1Gbit - 948 Mb/s if simplex and 1470 if duplex)
 - 1Gbit = 0.125 GB = 1000 Mb = 100 MB) but that is 100%

37



OTHER NETWORK

If 10Gb network check out Gareth's Webinar

- https://www.ibm.com/developerworks/wikis/download/attachments/153124943/7_PowerVM_10Gbit_Ethernet.pdf?version=1

netstat -v

- Look for overflows and memory allocation failures
 - Max Packets on S/W Transmit Queue: 884
 - S/W Transmit Queue Overflow: 9522
- "Software Xmit Q overflows" or "packets dropped due to memory allocation failure"
 - Increase adapter xmit queue
 - Use lsattr -EL ent? To see setting
- Look for receive errors or transmit errors
- dma underruns or overruns
- mbuf errors

lparstat 2

- Look for high vcswh – indicator that entitlement may be too low

tcp_nodelay (or tcp_nodelayack)

- Disabled by default
- 200ms delay by default as it waits to piggy back acks on packets

Also check errpt – people often forget this

38



ENTSTAT -V

ETHERNET STATISTICS (ent18) :
 Device Type: Shared Ethernet Adapter
 Elapsed Time: 44 days 4 hours 21 minutes 3 seconds

Transmit Statistics:	Receive Statistics:
-----	-----
Packets: 94747296468	Packets: 94747124969
Bytes: 99551035538979	Bytes: 99550991883196
Interrupts: 0	Interrupts: 22738616174
Transmit Errors: 0	Receive Errors: 0
Packets Dropped: 0	Packets Dropped: 286155
	Bad Packets: 0

Max Packets on S/W Transmit Queue: 712
 S/W Transmit Queue Overflow: 0
 Current S/W+H/W Transmit Queue Length: 50

Elapsed Time: 0 days 0 hours 0 minutes 0 seconds

Broadcast Packets: 3227715	Broadcast Packets: 3221586
Multicast Packets: 3394222	Multicast Packets: 3903090
No Carrier Sense: 0	CRC Errors: 0
DMA Underrun: 0	DMA Overrun: 0
Lost CTS Errors: 0	Alignment Errors: 0
Max Collision Errors: 0	No Resource Errors: 286155 check those tiny, etc Buffers
Late Collision Errors: 0	Receive Collision Errors: 0
Deferred: 0	Packet Too Short Errors: 0
SQE Test: 0	Packet Too Long Errors: 0
Timeout Errors: 0	Packets Discarded by Adapter: 0
Single Collision Count: 0	Receiver Start Count: 0
Multiple Collision Count: 0	

Current HW Transmit Queue Length: 50



39

ENTSTAT -V VIO

SEA

Transmit Statistics:	Receive Statistics:
-----	-----
Packets: 83329901816	Packets: 83491933633
Bytes: 87482716994025	Bytes: 87620268594031
Interrupts: 0	Interrupts: 18848013287
Transmit Errors: 0	Receive Errors: 0
Packets Dropped: 0	Packets Dropped: 67836309
	Bad Packets: 0

Max Packets on S/W Transmit Queue: 374
 S/W Transmit Queue Overflow: 0
 Current S/W+H/W Transmit Queue Length: 0

Elapsed Time: 0 days 0 hours 0 minutes 0 seconds

Broadcast Packets: 1077222	Broadcast Packets: 1075746
Multicast Packets: 3194318	Multicast Packets: 3194313
No Carrier Sense: 0	CRC Errors: 0
DMA Underrun: 0	DMA Overrun: 0
Lost CTS Errors: 0	Alignment Errors: 0
Max Collision Errors: 0	No Resource Errors: 67836309

Virtual I/O Ethernet Adapter (I-lan) Specific Statistics:

Hypervisor Send Failures: 4043136	"No Resource Errors" can occur when the appropriate amount of memory can not be added quickly to vent buffer space for a workload situation.
Receiver Failures: 4043136	
Send Errors: 0	

Hypervisor Receive Failures: 67836309



40

BUFFERS

Virtual Trunk Statistics

Receive Information

Receive Buffers

Buffer Type	Tiny	Small	Medium	Large	Huge
Min Buffers	512	512	128	24	24
Max Buffers	2048	2048	256	64	64
Allocated	513	2042	128	24	24
Registered	511	506	128	24	24
History					
Max Allocated	532	2048	128	24	24
Lowest Registered	502	354	128	24	24

“Max Allocated” represents the maximum number of buffers ever allocated
 “Min Buffers” is number of pre-allocated buffers
 “Max Buffers” is an absolute threshold for how many buffers can be allocated

```
chdev -l <veth> -a max_buf_small=4096 -P
chdev -l <veth> -a min_buf_small=2048 -P
```

Above increases min and max small buffers for the virtual ethernet adapter configured for the SEA above



41

Volume groups and file systems



42

BASICS

•Data layout will have more impact than most tunables

•Plan in advance

•Large hdisks are evil

- I/O performance is about bandwidth and reduced queuing, not size
- 10 x 50gb or 5 x 100gb hdisk are better than 1 x 500gb
- Also larger LUN sizes may mean larger PP sizes which is not great for lots of little filesystems
- Need to separate different kinds of data i.e. logs versus data

•The issue is queue_depth

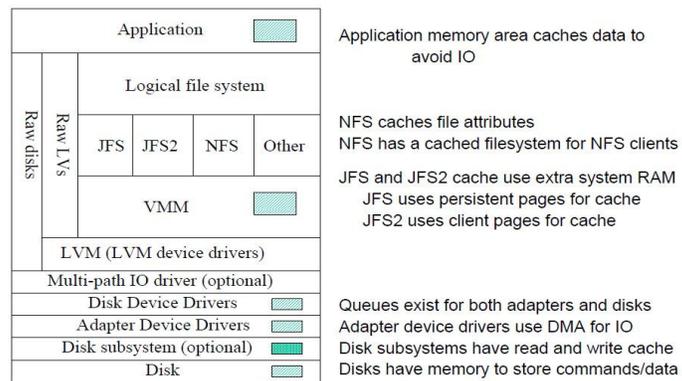
- In process and wait queues for hdisks
- In process queue contains up to queue_depth I/Os
- hdisk driver submits I/Os to the adapter driver
- Adapter driver also has in process and wait queues
- SDD and some other multi-path drivers will not submit more than queue_depth IOs to an hdisk which can affect performance
- Adapter driver submits I/Os to disk subsystem
- Default client qdepth for vSCSI is 3
 - chdev -l hdisk? -a queue_depth=20 (or some good value)
- Default client qdepth for NPIV is set by the Multipath driver in the client

43



From: PE23 Disk I/O Tuning in AIX v6.1 – Dan Braden and Steven Nasypany, October 2010

The AIX IO stack



Legend: ■ Write cache, ■ Read cache or memory area used for IO

IOs can be coalesced (good) or split up (bad) as they go thru the IO stack
 IOs adjacent in a file/LV/disk can be coalesced
 IOs greater than the maximum IO size supported will be split up



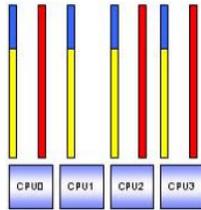
© 2010 IBM Corporation

44



IO WAIT AND WHY IT IS NOT NECESSARILY USEFUL

SMT2 example for simplicity



System has 3 threads blocked (red threads)
 SMT is turned on
 There are 4 threads ready to run so they get dispatched and each is using 80% user and 20% system

Metrics would show:

$$\%user = .8 * 4 / 4 = 80\%$$

$$\%sys = .2 * 4 / 4 = 20\%$$

Idle will be 0% as no core is waiting to run threads
 IO Wait will be 0% as no core is idle waiting for IO to complete as something else got dispatched to that core

SO we have IO wait
 BUT we don't see it

45



SAR -D

sar -d 2 6 shows:

	device	%busy	avque	r+w/s	Kbs/s	await	avserv
	hdisk7	0	0.0	2	160	0.0	1.9
	hdisk8	19	0.3	568	14337	23.5	2.3
avque	hdisk9	2	0.0	31	149	0.0	0.9

Average IOs in the wait queue

Waiting to get sent to the disk (the disk's queue is full)

Values > 0 indicate increasing queue_depth may help performance

Used to mean number of IOs in the disk queue

await

Time waiting in the wait queue (ms)

avserv

I/O service time when sent to disk (ms)

46



ADAPTER QUEUE PROBLEMS

Look at BBBF Tab in NMON Analyzer or run fcstat command

Adapter device drivers use DMA for IO

From **fcstat** on each fcs

NOTE these are since boot

FC SCSI Adapter Driver Information

No DMA Resource Count: 0

No Adapter Elements Count: 2567

No Command Resource Count: 34114051

- No DMA resource – adjust max_xfer_size
- No adapter elements – adjust num_cmd_elems
- No command resource - adjust num_cmd_elems

If using NPIV make changes to VIO and client, not just VIO

49



ADAPTER TUNING

fcs0

bus_intr_lvl	115	Bus interrupt level	False	
bus_io_addr	0xdfc00	Bus I/O address	False	
bus_mem_addr	0xe8040000	Bus memory address	False	
init_link	al	INIT Link flags	True	
intr_priority	3	Interrupt priority	False	
lg_term_dma	0x800000	Long term DMA	True	
max_xfer_size	0x100000	Maximum Transfer Size	True	(16MB DMA)
num_cmd_elems	200	Maximum number of COMMANDS to queue to the adapter	True	
pref_alpa	0x1	Preferred AL_PA	True	
sw_fc_class	2	FC Class for Fabric	True	

Changes I often make (test first)

max_xfer_size 0x200000 Maximum Transfer Size True **128MB DMA area for data I/O**
 num_cmd_elems 1024 Maximum number of COMMANDS to queue to the adapter True
lg_term_dma is the DMA area for control I/O

Check these are ok with your disk vendor!!!

```
chdev -l fcs0 -a max_xfer_size=0x200000 -a num_cmd_elems=1024 -P
chdev -l fcs1 -a max_xfer_size=0x200000 -a num_cmd_elems=1024 -P
```

Remember make changes too both VIO servers and client LPARs if using NPIV
 VIO server setting must be at least as large as the client setting

50



MY VIO SERVER AND NPIV CLIENT ADAPTER SETTINGS

VIO SERVER
 #lsattr -El fcs0
 lg_term_dma 0x800000 Long term DMA True
 max_xfer_size 0x200000 Maximum Transfer Size True
 num_cmd_elems 1024 Maximum number of COMMANDS to queue to the
 adapter True

NPIV Client (running at defaults before changes)
 #lsattr -El fcs0
 lg_term_dma 0x800000 Long term DMA True
 max_xfer_size 0x200000 Maximum Transfer Size True
 num_cmd_elems 512 Maximum Number of COMMAND Elements True



PARAMETER SETTINGS - SUMMARY

PARAMETER	DEFAULTS			NEW SET ALL TO	
	AIXv5.3	AIXv6	AIXv7		
NETWORK (no)					
rfc1323	0	0	0	1	
tcp_sendspace	16384	16384	16384	262144 (1Gb)	
tcp_recvspace	16384	16384	16384	262144 (1Gb)	
udp_sendspace	9216	9216	9216	65536	
udp_recvspace	42080	42080	42080	655360	
MEMORY (vmo)					
minperm%	20	3	3	3	
maxperm%	80	90	90	90	JFS, NFS, VxFS, JFS2
maxclient%	80	90	90	90	JFS2, NFS
lru_file_repage	1	0	0	0	
lru_poll_interval	?	10	10	10	
Minfree	960	960	960	calculation	
Maxfree	1088	1088	1088	calculation	
page_steal_method	0	0 / 1 (TL)	1	1	
JFS2 (ioo)					
j2_maxPageReadAhead	128	128	128	as needed	
j2_dynamicBufferPreallocation	16	16	16	as needed	



OTHER INTERESTING TUNABLES

noatime

- Why write a record every time you read or touch a file?
- mount command option
- Use for redo and archive logs

Release behind (or throw data out of file system cache)

- rbr – release behind on read
- rbw – release behind on write
- rbrw – both

Read the various AIX Difference Guides:

- <http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=aix+AND+differences+AND+guide>

53



ADAPTER PRIORITIES AFFECT PERFORMANCE

Power 770 Layout		9117-MMC												
CEC	Top	123456 has GX cables				Bottom	2468ab				5877 pcie only I/O Drawer 123487			
	Slot	Desc	Pri	Alloc	Slot	Desc	Pri	Alloc	Slot	Desc	Pri	Alloc	IOC	
	C1	8GB DP fibre	1	lpar1	C1	8GB DP fibre	1	lpar1	C1	8GB DP fibre	1	vio1	1	
	C2	4PT 10/100/1000	3	lpar1	C2	4PT 10/100/1000	3	lpar1	C2	4PT 10/100/1000	3		1	
	C3	8GB DP fibre	5	vio2	C3	8GB DP fibre	5	vio1	C3		5		1	
	C4	4PT 10/100/1000	6	vio2	C4	4PT 10/100/1000	6	vio1	C4	8GB DP fibre	2	vio2	2	
	C5	8GB DP fibre	2	vio1	C5	8GB DP fibre	2	vio2	C5	4PT 10/100/1000	4		2	
	C6	4PT 10/100/1000	4	vio1	C6	4PT 10/100/1000	4	vio2	C6	4GB DP fibre	6	lpar1	2	
									C7	4GB DP fibre	7		3	
	D1	146GB disk		vio1	D1	146GB disk		vio1	C8		8		3	
	D4	146GB disk		vio2	D4	146GB disk		vio2	C9		9		3	
									C10		10		3	

54



I/O BANDWIDTH

PCIe2 LP 8Gb 4 port Fibre HBA

- Data throughput 3200 MB/ps FDX per port
- IOPS 200,000 per port
- <http://www.redbooks.ibm.com/technotes/tips0883.pdf>
- Can run at 2Gb, 4Gb or 8Gb

PCIe2 8Gb 1 or 2 port Fibre HBA

- Data throughput 1600 MB/s FDX per port
- IOPS Up to 200,000 per port

Above are approximate taken from card specifications
 Look at DISK_SUMM tab in nmon analyzer
 Sum reads and writes, figure out the average and max
 Then divide by 1024 to get MB/s

55



ORACLE Asynchronous I/O and Concurrent I/O

56



ASYNCR I/O - v5.3

Total number of AIOs in use

pstat -a | grep aios | wc -l
 Maximum AIO servers started since boot
 servers per cpu True

NB - maxservers is a per processor setting in AIX 5.3

Or new way for Posix AIOs is:

ps -k | grep aio | wc -l
 4205

At AIX v5.3 tl05 this is controlled by aioo command

Also iostat -A

THIS ALL CHANGES IN AIX V6 - SETTINGS WILL BE UNDER IOO THERE

lsattr -El aio0

autoconfig	defined STATE to be configured at system restart	True	
fastpath	enable State of fast path		True
kproprio	39 Server PRIORITY		True
maxreqs	4096 Maximum number of REQUESTS	True	
maxservers	10 MAXIMUM number of servers per cpu	True	
minservers	1 MINIMUM number of servers	True	

AIO is used to improve performance for I/O to raw LVs as well as filesystems.

57



IOSTAT -A

iostat -A async IO

System configuration: lcpu=16 drives=15

aio: avgc avfc maxg maif maxr avg-cpu: % user % sys % idle % iowait

150	0	5652	0	12288		21.4	3.3	64.7	10.6
-----	---	------	---	-------	--	------	-----	------	------

Disks: % tm_act Kbps tps Kb_read Kb_wrtn

hdisk6	23.4	1846.1	195.2	381485298	61892856
hdisk5	15.2	1387.4	143.8	304880506	28324064
hdisk9	13.9	1695.9	163.3	373163558	34144512

If maxg close to maxr or maxservers then increase maxreqs or maxservers

Old calculation - no longer recommended

minservers = active number of CPUs or 10 whichever is the smaller number
 maxservers = number of disks times 10 divided by the active number of CPUs
 maxreqs = 4 times the number of disks times the queue depth

***Reboot anytime the AIO Server parameters are changed

58



ASYNCR I/O – AIX V6 AND V7

No more smit panels and no AIO servers start at boot
 Kernel extensions loaded at boot
 AIO servers go away if no activity for 300 seconds
 Only need to tune maxreqs normally

ioo -a -F | more

```
aio_active = 0
aio_maxreqs = 65536
aio_maxservers = 30
aio_minservers = 3
aio_server_inactivity = 300
posix_aio_active = 0
posix_aio_maxreqs = 65536
posix_aio_maxservers = 30
posix_aio_minservers = 3
posix_aio_server_inactivity = 300
```

pstat -a | grep aio

```
22 a 1608e 1 1608e 0 0 1 aioPpool
24 a 1804a 1 1804a 0 0 1 aioLpool
```

##Restricted tunables

```
aio_fastpath = 1
aio_fsfastpath = 1
aio_kprocprio = 39
aio_multitidsusp = 1
aio_sample_rate = 5
aio_samples_per_cycle = 6
posix_aio_fastpath = 1
posix_aio_fsfastpath = 1
posix_aio_kprocprio = 39
posix_aio_sample_rate = 5
posix_aio_samples_per_cycle = 6
```



AIO RECOMMENDATIONS

Oracle now recommending the following as starting points

	5.3	6.1 or 7 (non CIO)
minservers	100	3 - default
maxservers	200	200
maxreqs	16384	65536 – default

These are per CPU

So for lcpu=10 and maxservers=100 you get 1000 aioservers

AIO applies to both raw I/O and file systems

Grow maxservers as you need to



DIO AND CIO

DIO

- Direct I/O
- Around since AIX v5.1, also in Linux
- Used with JFS
- CIO is built on it
- Effectively bypasses filesystem caching to bring data directly into application buffers
- Does not like compressed JFS or BF (lfe) filesystems
 - Performance will suffer due to requirement for 128kb I/O (after 4MB)
- Reduces CPU and eliminates overhead copying data twice
- Reads are asynchronous
- No filesystem readahead
- No lrucl or syncd overhead
- No double buffering of data
- Inode locks still used
- Benefits heavily random access workloads

61



DIO AND CIO

CIO

- Concurrent I/O – AIX only, not in Linux
- Only available in JFS2
- Allows performance close to raw devices
- **Designed for apps (such as RDBs) that enforce write serialization at the app**
- Allows non-use of inode locks
- Implies DIO as well
- Benefits heavy update workloads
- Speeds up writes significantly
- Saves memory and CPU for double copies
- No filesystem readahead
- No lrucl or syncd overhead
- No double buffering of data
- **Not all apps benefit from CIO and DIO – some are better with filesystem caching and some are safer that way**

When to use it

- Database DBF files, redo logs and control files and flashback log files.
- Not for Oracle binaries or archive log files

62



DIO/CIO ORACLE SPECIFICS

Use CIO where it will benefit you

- Do not use for Oracle binaries
- Ensure redo logs and control files are in their own filesystems with the correct (512) blocksize
 - Use `lsfs -q` to check blocksizes
- I give each instance its own filesystem and their redo logs are also separate

Leave `DISK_ASYNC_IO=TRUE` in Oracle
Tweak the maxservers AIO settings

Remember CIO uses DIO under the covers

If using JFS

- Do not allocate JFS with BF (LFE)
- It increases DIO transfer size from 4k to 128k
- 2gb is largest file size
- Do not use compressed JFS – defeats DIO

63



TELLING ORACLE TO USE CIO AND AIO

If your Oracle version (10g/11g) supports it then configure it this way:

There is no default set in Oracle 10g do you need to set it

Configure Oracle Instance to use CIO and AIO in the init.ora (PFILE/SPFILE)

```
disk_async_io      = true      (init.ora)
```

```
filesystemio_options = setall  (init.ora)
```

Note if you do backups using system commands while the database is up then you will need to use the 9i method below for v10 or v11

If not (i.e. 9i) then you will have to set the filesystem to use CIO in the /etc filesystems

```
options            = cio      (/etc/filesystems)
```

```
disk_async_io      = true     (init.ora)
```

Do not put anything in the filesystem that the Database does not manage – remember there is no inode lock on writes

Or you can use ASM and let it manage all the disk automatically

Also read Metalink Notes #257338.1, #360287.1

See Metalink Note 960055.1 for recommendations

Do not set it in both places (config file and /etc/filesystems)

64



DEMOTED I/O IN ORACLE

CIO write fails because IO is not aligned to FS blocksize

- i.e app writing 512 byte blocks but FS has 4096

Ends up getting redone

- Demoted I/O consumes more kernel CPU
- And more physical I/O

To find demoted I/O (if JFS2)

```
trace -aj 59B,59C ; sleep 2 ; trcstop ; trcrpt -o directio.trcrpt
grep -i demoted directio.trcrpt
```

Look in the report for:

```
JFS2 IO dio demoted:
1000 1000 dio demoted:
```

65



USEFUL LINKS

AIX Wiki

- <https://www.ibm.com/developerworks/wikis/display/WikiPtype/AIX>

HMC Scanner

- <http://www.ibm.com/developerworks/wikis/display/WikiPtype/HMC+Scanner>

Workload Estimator

- <http://ibm.com/systems/support/tools/estimator>

Performance Tools Wiki

- <http://www.ibm.com/developerworks/wikis/display/WikiPtype/Performance+Monitoring+Tools>

Performance Monitoring

- <https://www.ibm.com/developerworks/wikis/display/WikiPtype/Performance+Monitoring+Documentation>

Other Performance Tools

- <https://www.ibm.com/developerworks/wikis/display/WikiPtype/Other+Performance+Tools>
- Includes new advisors for Java, VIOS, Virtualization

VIOS Advisor

- <https://www.ibm.com/developerworks/wikis/display/WikiPtype/Other+Performance+Tools#OtherPerformanceTools-VIOSPA>

66



REFERENCES

Simultaneous Multi-Threading on POWER7 Processors by Mark Funk

- http://www.ibm.com/systems/resources/pwrsysperf_SMT4OnP7.pdf

Processor Utilization in AIX by Saravanan Devendran

- <https://www.ibm.com/developerworks/mydeveloperworks/wikis/home?lang=en#/wiki/Power%20Systems/page/Understanding%20CPU%20utilization%20on%20AIX>

Rosa Davidson Back to Basics Part 1 and 2 –Jan 24 and 31, 2013

- <https://www.ibm.com/developerworks/mydeveloperworks/wikis/home?lang=en#/wiki/Power%20Systems/page/AIX%20Virtual%20User%20Group%20-%20USA>

Nigel – PowerVM User Group

- <https://www.ibm.com/developerworks/mydeveloperworks/wikis/home?lang=en#/wiki/Power%20Systems/page/PowerVM%20technical%20webinar%20series%20on%20Power%20Systems%20Virtualization%20from%20IBM%20web>

SG24-7940 - PowerVM Virtualization - Introduction and Configuration

- <http://www.redbooks.ibm.com/redbooks/pdfs/sg247940.pdf>

SG24-7590 – PowerVM Virtualization – Managing and Monitoring

- <http://www.redbooks.ibm.com/redbooks/pdfs/sg247590.pdf>

SG24-8080 – Power Systems Performance Guide – Implementing and Optimizing

- <http://www.redbooks.ibm.com/redbooks/pdfs/sg248080.pdf>

SG24-8079 – Power 7 and 7+ Optimization and Tuning Guide

- <http://www.redbooks.ibm.com/redbooks/pdfs/sg248079.pdf>

Redbook Tip on Maximizing the Value of P7 and P7+ through Tuning and Optimization

- <http://www.redbooks.ibm.com/technotes/tips0956.pdf>

TIPS TO KEEP OUT OF TROUBLE

Check the performance apars have all been installed

- Yes this means you need to stay current
- See Stephen Nasypany and Rosa Davidson Optimization Presentations
- Keep firmware up to date
- In particular, look at the firmware history for your server to see if there are performance problems fixed

Information on the firmware updates can be found at:

- <http://www-933.ibm.com/support/fixcentral/>

Firmware history including release dates can be found at:

- Power7 Midrange
 - <http://download.boulder.ibm.com/ibmdl/pub/software/server/firmware/AM-Firmware-Hist.html>
- Power7 High end
 - <http://download.boulder.ibm.com/ibmdl/pub/software/server/firmware/AL-Firmware-Hist.html>
- Ensure software stack is current
- Ensure compilers are current and that compiled code turns on optimization
- To get true MPIO run the correct multipath software
- Ensure system is properly architected (VPs, memory, entitlement, etc)
- DOCUMENTATION

THANK YOU FOR YOUR TIME



If you have questions please email me at:
lynchj@forsythe.com

This presentation at:
<http://www.circle4.com/papers/common-performance.pdf>

69



BACKUP SLIDES PERFORMANCE TOOLS

70



TOOLS

topas

- New -L flag for LPAR view

nmon

nmon analyzer

- Windows tool so need to copy the .nmon file over in ascii mode
 - Opens as an excel spreadsheet and then analyses the data
 - Also look at nmon consolidator
- ### sar
- sar -A -o filename 2 30 >/dev/null
 - Creates a snapshot to a file – in this case 30 snaps 2 seconds apart
 - Must be post processed on same level of system

errpt

Check for changes from defaults

<https://www.ibm.com/developerworks/wikis/display/WikiPtype/Other+Performance+Tools>

ioo, vmo, schedo, vmstat -v

lvmo

lparstat, mpstat

iostat

Check out Alphaworks for the Graphical LPAR tool

Ganglia - <http://ganglia.info>

Nmonrrd and nmon2web and pGraph

Commercial IBM

- PM for AIX
- Performance Toolbox
- Tivoli ITM

71



OTHER TOOLS

filemon

- filemon -v -o filename -O all
- sleep 30
- trcstop

pstat to check async I/O in 5.3

- pstat -a | grep aio | wc -l

perfpmr to build performance info for IBM if reporting a PMR

- /usr/bin/perfpmr.sh 300

72



NMON

`nmon -ft -A -s 15 -c 120`

- Grabs a 30 minute nmon snapshot with async I/O

`nmon -ft -A -M -L -^ -s 15 -c 120`

- Same as above but includes large pages and some other features

Must be running nmon12e or higher

Nmon comes with AIX at 5.3 tl09 or 6.1 tl01 and higher BUT on 5.3 I download the latest version from the web so I get the latest v12 for sure

Creates a file in the working directory that ends .nmon

This file can be transferred to your PC and interpreted using nmon analyser or other tools

`nmon -f -O` - now gets seastats for VIO server

`nmon -f -K` - dump libperfstat structures

<http://www.ibm.com/developerworks/wikis/display/WikiPtype/nmon>

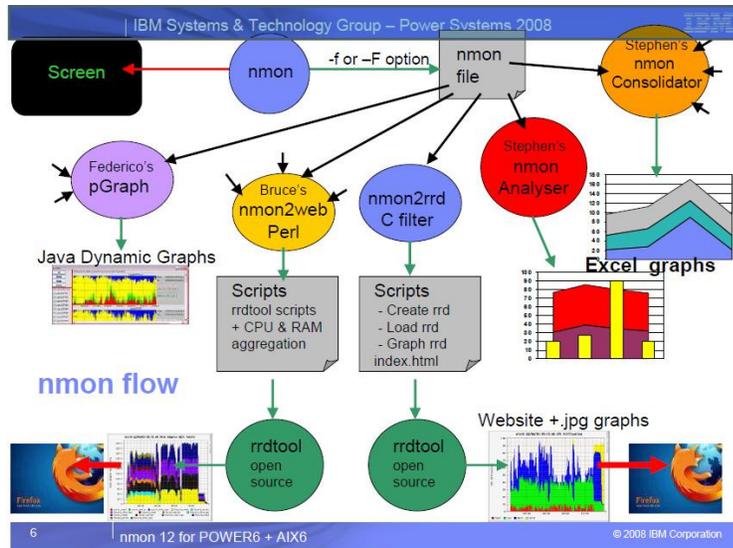
<http://www.ibm.com/developerworks/wikis/display/WikiPtype/nmonanalyser>

<http://www.ibm.com/developerworks/wikis/display/WikiPtype/nmonconsolidator>



73

NMON



Courtesy Nigel Griffiths - IBM

74



NMON ON POWER6 & AIX6 + - NEW FEATURES FOR V12

- Disk Service Times
- Selecting Particular Disks
- Time Drift
- Multiple Page Sizes
- Timestamps in UTC & no. of digits
- More Kernel & Hypervisor Stats *
- High Priority nmon
- Advanced, POWER6 and AIX6 items
- Virtual I/O Server SEA
- Partition Mobility (POWER6)
- WPAR & Application Mobility (AIX6)
- Dedicated Donating (POWER6)
- Folded CPU count (SPLPAR)
- Multiple Shared Pools (POWER6)
- Fibre Channel stats via entstat

75



PERFORMANCE WIKI

Overview New to Forums Wikis

Other Performance Tools

View Attachments (39) Info

Browse Space

Added by nag, last edited by nag on Feb 14, 2012 (view change)
Labels: (None)

Performance - Other Tools (non AIX commands)



Quick Links to this page:

NEW: [Java Performance Advisor](#) · [VIOS Performance Advisor](#) · [Virtualization Performance Advisor](#)
nmon based: [nmon](#) · [nmon-Analyser](#) · [nmon-Consolidator](#) · [pGraph](#) · [topas-CEC-nmon-Analyser](#) · [nmon2web](#) · [nmon2rrd](#)
Regular: [gmon v71n](#) · [Visual Performance Analyzer](#) · [Roll-Your-Own](#) · [nstress-Tools](#)
LPAR-Monitor · [SEA-Monitor](#) · [HMC-LPAR-data-2-rrd](#) · [WLM](#) · [rPerf](#) · [rvaacct](#) · [nworms](#)
Products and larger tools: [PM](#) · [Ganglia](#) · [Munin](#) · [Galileo](#) · [Performance-Toolbox](#)

VIOS Performance Advisor



The VIOS advisor is an application that runs within the customer's VIOS for a user specified amount of time (hours), which polls and collects key performance metrics before analyzing results and providing a health check report and proposes changes to the environment or areas to investigate further. The goal of the VIOS advisor is not to provide another monitoring tool, but instead have an expert system view performance metrics already available to the customer and make assessments and recommendations based on the expertise and experience available within the IBM systems performance group.
 Take this link to the [VIOS Performance Advisor](#)

76



VIOS ADVISOR

<https://www.ibm.com/developerworks/wikis/display/WikiPtype/VIOS+Advisor>
 Application that collects performance metrics and does a health check
 Productized in VIOS 2.2.2
 Following slides run on a production VIO during a regular production day



77

VIOS ADVISOR

The ratings and recommendations in the table below were chosen with the following information:

Hostname: vio1. .com
PartitionID: 2
Monitoring Start Time: 03/09 11:45:19
Monitoring Stop Time: 03/09 13:45:19 **Duration:** 120 min
IBM Systems Workload Estimator link: <http://ibm.com/systems/support/tools/estimator> (VIOS Sizings)

SYSTEM - CONFIGURATION		
	Name	Value
	Processor Family	POWER6
	Server Model	IBM,9117-MMA
	Server Frequency	4.208 GHz
	Server - Online CPUs	10 cores
	Server - Maximum Supported CPUs	16 cores
	VIOS Level	2.2.0.13-FP24 SP-03
	VIOS Advisor Release	121211B



78

VIOS ADVISOR

VIOS - CPU							
	Name	Measured Value	Recommended Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowes 5=highes
	CPU Capacity	1.0 ent	-	03/09 11:45:19	-	n/a	n/a
	CPU Consumption	avg:5.4% (cores:0.1) high:40.2% (cores:0.5)	-	-	-	n/a	n/a
	Processing Mode	Shared CPU, (UnCapped)	-	03/09 11:45:19	-	n/a	n/a
	Variable Capacity Weight	200	-	03/09 11:45:19	-	n/a	n/a
	Virtual Processors	2 vCPUs	-	03/09 11:45:19	-	n/a	n/a
	SMT Mode	SMT2	-	03/09 11:45:19	-	n/a	n/a

FORSYTHE

VIOS ADVISOR

SYSTEM - SHARED PROCESSING POOL							
	Name	Measured Value	Recommended Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowes 5=highes
	Shared Pool Monitoring	enabled	-	03/09 11:45:19	-	n/a	n/a
	Shared Processing Pool Capacity	10.0 ent.	-	03/09 11:45:19	-	n/a	n/a
	Free CPU Capacity	avg_free:9.4 ent. lowest_free:7.7 ent.	-	-	-	n/a	n/a

FORSYTHE

VIOS ADVISOR

VIOS - I/O ACTIVITY

	Name	Value
	Disk I/O Activity	avg: 229 iops @ 32KB peak: 1916 iops @ 137KB
	Network I/O Activity	[avgSend: 0 iops 0.0MBps , avgRcv: 0 iops 0.0MBps] [peakSend: 0 iops 0.0MBps , peakRcv: 0 iops 0.0MBps]

81


VIOS ADVISOR

VIOS - DISK ADAPTERS

	Name	Measured Value	Recommended Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowest 5=highest
	FC Adapter Count	3	-	03/09 11:45:19	-	n/a	n/a
	FC Avg IOps	avg: 77 iops @ 32KB	-	03/09 11:45:19	03/09 13:45:19	n/a	n/a
	FC Adapter Utilization	optimal	-	-	-	n/a	n/a
	FC Port Speeds	running at speed	-	-	-	n/a	n/a

82


VIOS ADVISOR

VIOS - DISK DRIVES							
	Name	Measured Value	Recommended Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowest 5=highest
	Physical Drive Count	93	-	03/09 11:45:19	-	n/a	n/a
	I/Os Blocked	optimal	-	-	-	n/a	n/a
	Long I/O Latency (hdisk3)	avg:9.7ms (9.7 + 0.0) high:11.5ms (11.5 + 0.0)	Range: 8-12ms	03/09 12:35:58	03/09 13:44:02	n/a	n/a

83

VIOS ADVISOR

VIOS - MEMORY							
	Name	Measured Value	Recommended Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowest 5=highest
	Real Memory	4.000 GB	7.000 GB	03/09 11:45:19	-	1	2
	Available Memory	0.346 GB	1.5 GB Avail.	03/09 11:45:39	03/09 13:45:05	n/a	n/a
	Paging Rate	0.2 MB/s pg rate	-	-	-	n/a	n/a
	Paging Space Size	8.000 GB	-	03/09 11:45:19	-	n/a	n/a
	Free Paging Space	7.923 GB free	-	-	-	n/a	n/a
	Pinned Memory	1.262 GB pinned	-	-	-	n/a	n/a

84