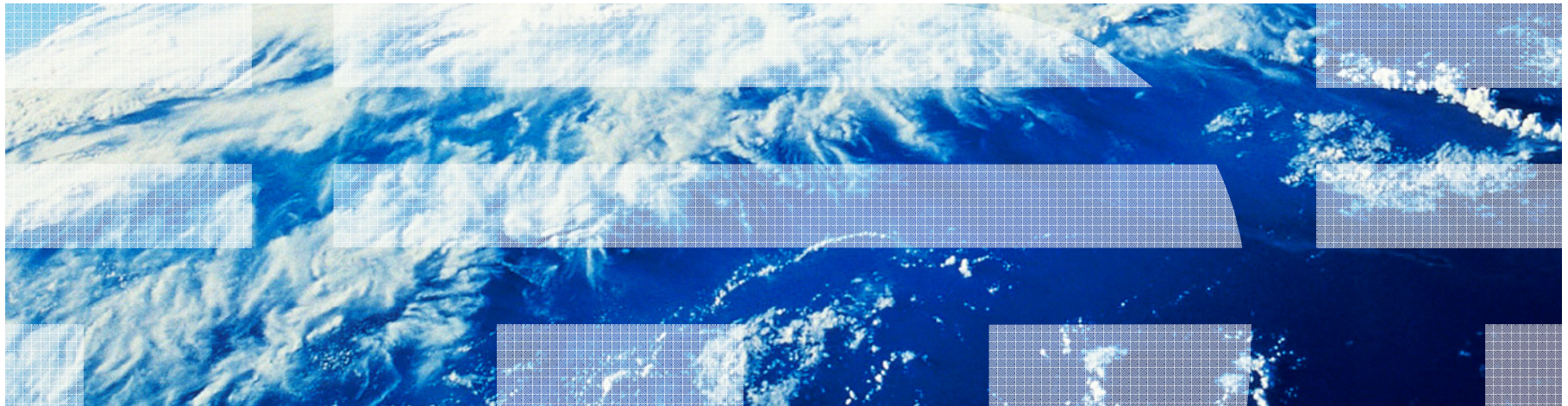


pHS10 Best Practices to Optimize AIX on POWER7

Grover Davidson - Development Support



IBMTECHU.COM

Sign out

Home

Update profile

Message board

All	Monday	Tuesday	Wednesday	Thursday	Friday
8:00-8:50	9:00-10:15	10:30-11:45	1:00-2:15	2:30-3:45	4:15-5:30

Tracks (All) ▼

Evaluations

Keynote Eval

Session Evals

Overall Conference

Eval Summary

Your personal agenda plan will appear to the right of this section, when you first add a session. Each session in your plan is a hyperlink to the session evaluation page.

1. Use the section above to filter session data. Results will appear here. For example, selecting Wednesday, search: alex, will display all sessions on Wednesday containing text "alex".
2. To add session to your personal agenda plan, press the blue button next to the appropriate session slot.
3. To remove a session from your personal agenda plan, press the green button next to the appropriate session slot.

Planning

Agenda Planner

Your Agenda

Daily Changes

Downloads

Access material

Survey

Marketing Survey

IBMTECHU.COM

- **IBM STG Technical Universities & Conferences web portal**

- **Direct link: ibmtechu.com/uk**

- **KEY FEATURES...**
 - Create a personal agenda using the agenda planner
 - View the agenda and agenda changes
 - Use the agenda search to find the sessions and/or
 - Download presentations
 - **Submit Session and Conference Evaluations**



Factors influencing AIX performance on POWER7

- Performance analyzers
- White papers & FAQs
- Firmware levels
- LPAR placement
- AIX performance fixes
- LPAR Entitlement
- Libc malloc fixes
- DMA Overruns
- AIX netcd
- AIX Tunables
- VLAN Tuning
- 64K pages

Performance Advisors

- VIO Servers

- <http://www.ibm.com/developerworks/wikis/display/WikiPtype/VIOS+Advisor>

- LPAR Performance Advisor

- <https://www.ibm.com/developerworks/wikis/display/WikiPtype/PowerVM+Virtualization+performance+lpar+advisor>

- Java Performance Advisor

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

White Papers and FAQs

- Power7 Virtualization Best Practice Guide
 - https://www.ibm.com/developerworks/wikis/download/attachments/53871915/P7_virtualization_bestpractice.doc?version=1
- Java Performance on POWER7 – Best Practice
 - http://www-304.ibm.com/partnerworld/wps/servlet/ContentHandler/stg_ast_sys_java_performance_on_power7
- Oracle Architecture and Tuning on AIX v2.20
 - <http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP100883>

Firmware levels for 770/780/795

- Minimum level should be 730_066
 - Below this level has a number of significant issue that affect performance
 - Many are very difficult to detect and diagnose
 - Any system below this level should consider updating a priority
- Preferred level currently available is 730-078
 - A problem was fixed that impacted performance if profiling was enabled in one or more partitions. Performance profiling is enabled:
 - In an AIX or VIOS partition using the tprof (-a, -b, -B, -E option) command or pmctl (-a, -E option) command.
 - Tprof is a profiling command used to diagnose performance issue
 - Event based profiling can result in performance issues that require the CEC to be rebooted
- ALWAYS check for new levels and what issues they address when a maintenance window exists where an update can be performed
- More likely to see issues due to hardware configuration

LPAR Placement

- LPAR placement can affect performance
- DLPAR operations increase probability of resource fragmentation
- See section 1.3 of the “POWER7 Virtualization Best Practice Guide” if the if placement is suspected to be part of the problem
- Follow the directions at the end of section 1.3.2 to help optimize the placement of resources

AIX Performance Bundles

- There are a number of performance related fixes available for AIX
- They address serious and common issues
- A special service release was done to make them available to customers
- Recommended APARs/fileset levels are:

- IV12132 U846733 shipped 7100-01-03-1207 bos.mp64 7.1.1.3
- IV14644 U846561 shipped bos.mp64 7.1.0.18
- IV11778 U846207 shipped 6100-07-03-1207 bos.mp64 6.1.7.3
- IV13560 U849131 shipped bos.mp64 6.1.6.18
- IV13474 U842593 shipped bos.mp64 6.1.5.10

Fixes addressed

- address space lock contention
- SRAD load balancing issues on shared LPARs
- waitproc idle looping consumes cpu
- miscellaneous dispatcher/scheduling performance fixes
- crashed@sysmt+0001C0 – processor folding issue
- high PSMD threads CPU consumption
- paging 4K pages because of lack of 64K page demotions
- paging to page space on properly tuned system

Entitlement issues

- Use the Performance Advisors help you properly entitle the LPAR
- Attempt to avoid an LPAR running constantly over 100% entitlement
- Do not configure more VPs in a single LPAR than exist in the shared pool
- AIX is designed to maximize throughput and minimize response times
- As a result AIX will use maximum resources to process workload
- Right sizing the LPAR is key to effective use of resources

libc malloc issues

- Malloc is not correctly allocating memory from multiple heaps with a 32 bit application
- Symptom is heavy single lock contention within the application space when heavy malloc and free activity is occurring
- Does NOT affect 64 bit applications
- First found in a webserver showing performance issue issues
- APAR Description is MALLOC() IGNORES MULTIHEAPS IN 32-BIT MODE
- Install one of the following:
 - IV15057 U840899 unshipped bos.rte.libc 7.1.1.15
 - IV15066 U842969 unshipped bos.rte.libc 6.1.7.15
 - IV10131 U849882 unshipped bos.rte.libc 6.1.6.18
 - IV10058 U851309 unshipped bos.rte.libc 6.1.5.9

DMA overruns on FC/Ethernet adapters

- If the system has a combined FC/Ethernet adapter card one of the following fixes should be installed:
 - IV02894 U843415 devices.pciex.771000801410b003.rte 7.1.1.15
 - IV02894 U842413 devices.pciex.771000801410b003.rte 7.1.1.0
 - IV05880 U839066 devices.pciex.771000801410b003.rte
 - IV02568 U843033 devices.pciex.771000801410b003.rte 6.1.7.15
 - IV02568 U840951 devices.pciex.771000801410b003.rte 6.1.7.0
 - IV04974 U843166 devices.pciex.771000801410b003.rte 6.1.6.16
 - IV05267 U845251 devices.pciex.771000801410b003.rte 6.1.5.3
 - IV05673 U845188 devices.pciex.771000801410b003.rte 6.1.4.6
 - IV04302 U844681 devices.pciex.771000801410b003.rte 5.3.12.3
 - IV04417 U844844 devices.pciex.771000801410b003.rte 5.3.11.6
- Ignore the APAR description – it is misleading
- Checking to see if this should be installed:
 - lsdev –C | egrep –i 771000801410b003

No output means this adapter is not installed on the system

AIX 6.1 added a network cache daemon

- Called netcd and controlled via System Resource Controller
- Started with `'startsrc -s netcd'`
- Can act as a caching Name resolver
- Prevents repeated lookups to DNS for the same name
- Removes latency from overloaded DNS servers
- See config file man pages for further details
–`man netcd.conf`

AIX Tunables

- Do **NOT** carry forward tunings from and AIX 5.3 system to AIX 6.1
- Restricted tunables should **NOT** be changed on AIX 6.1/7.1 without advice from IBM support
- IO pacing should NOT be changed with strong justification and low values (like 24/33) can result in serious IO performance issues
- The sys0 attribute iostat should be set to false – it causes extra work to update disk usage data since boot time
- Disabling use_isno for the network may result in serious performance issues
- Remember to tune rfc1323/tcp_nodelay/tcp_sendspace/tcp_recvspace on a PER ADAPTER basis
- HEA/IVE adapter should have the rx_clsc set to 10G for low latency applications

Virtual Ethernet buffer tuning

- Check the buffer allocations on both the VIO Server and Client:

```
entstat -d entX
```

```
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
History					
Max Allocated	512	2048	128	24	24

- If Max Allocated equals Max Buffers then increase max buffers for that type
- Small buffers need to be increased in the example above:

```
chdev -l entX -a max_buf_small=4096 -a min_buf_small=2048 -P
```

- Requires a reboot to take affect
- For HEAVY traffic interfaces increase min to max value to eliminate dynamic buffer management
- CAUTION:** Setting these values too high with a large number of adapters can make the system fail to boot – Tune ONLY when needed

Network Tuning

- Enable watchdog threads with ifconfig
 - MUST be done each time rebooted
 - Enabled with: `ifconfig enX thread`
 - Disabled with: `ifconfig enX -thread`
- Enable `large_send` and `large_recv` via `chdev` and `ifconfig` on all VIO Servers and Clients:
 - `chdev -l entX -a large_send=true -a large_recv=true`
 - Enabled with: `ifconfig enX largesend`
 - Disabled with: `ifconfig enX -largesend`
 - No option for `largerecv` via `ifconfig`
 - See `ifconfig` man page for further details
- Systems running with SEA adapters should have `ifix` for IV07193 installed:
 - SEA thread lock contention prevents scale up
- Tuning should be done on a per-adapter basis
 - Use `'lsattr -El enX'` to see what attributes on the adapter can be tuned

When to consider 64k pages

- AIX performs dynamic page promotion/demotion between 4k pages to 64k pages beginning with AIX 6.1
- When 16 consecutive 4k pages are assigned VMM attempts to make them into a single 64k page
- This can significantly improve application performance and minimize memory usage
- Can be problematic for long running programs that fork a lot of child processes
- Consider altering the executable environment to use 64k pages via LDR_CNTRL:
LDR_CNTRL=TEXTFSIZE=64K@DATAFSIZE=64K@STACKFSIZE=64K@SH
MFSIZE=64K command_name
- Do NOT put this in /etc/environment
- Will use more memory if used incorrectly
- Symptoms include psmd consuming a lot of time and a high fork rate in SAR data

Summary

- Many resources including Performance Advisors and Whitepapers/FAQs exist
- Firmware and AIX fixes are critical to avoiding problems
- Entitlement and LPAR sizing is critical to proper performance and utilization
- netcd can improve latency by caching frequent resource lookups
- AIX 5.3 tunings do not apply to AIX 6.1/7.1
- VLAN tuning can have a significant affect
- 64K pages can improve performance