Cheryl's Hot Flashes

Cheryl Watson

Session 2543, August 27, 1999

Watson & Walker, Inc.
   publishers of 'Cheryl Watson's TUNING Letter'
http://www.watsonwalker.com

Technology ▪ Connections ▪ Results

Cheryl's Hot Flashes

- Virtual Storage Constraint
- Reporting CPU Busy
- Warnings
- WLM Quickstart Policy
- WLM
- MQSeries
- Web Sites
- APARs of Interest
- Oldies but Goodies
- HFS in R7

Virtual Storage Constraint

- Many upgrades to OS/390 reduce virtual storage private region below 16MB

    - Notice! There are some OS/390 R5+ sites who can provide an 11MB private region - don't penalize your users with an 8 MB region!

    - Primary reason for common storage growth is additional 24-bit LPA modules

    - Ways to decrease constraint:

        - Move UCBs above the the line (LOCANY)

        - Look for APARs relating to CSA and SQA

        - **<u>Remove modules from LPA (e.g. ISAM, unused ISPF modules)</u>**

        - See ISPF Planning manual - SC28-1298

        - Move buffers above the line

        - [Ref: CWTL, 1999, No. 3]

Reporting CPU Busy

● Many reports today use AVERAGE CPU BUSY%

CPU 0	32%
CPU 1	30%
CPU 2	28%
AVG	30%

● BUT - some sites vary CPs offline; or use dynamic ICFs

CPU 0	0%
CPU 1	46%
CPU 2	44%
AVG	45%

● SO, both are using the same amount of CPU, but the average CPU busy differs

Reporting CPU Busy

- Better technique - use Total CPU Busy of machine

CPU 0	32%
CPU 1	30%
CPU 2	28%
Total	90%

- When one not utilized

CPU 0	0%
CPU 1	46%
CPU 2	44%
Total	90%

- By using total CPU instead of average CPU, you can show utilization even though configuration changes

Warnings

● **From Lin Merritt**

   ■ **Invalid DB2 SRB time in DB2 101 records**

     – The 101 records provide TCB and SRB time by transaction

     – In a parallel environment (where DB2 is parallelizing a query), the TCB time for each piece is correct, but the SRB time for all the pieces are placed in each transaction

     – Thus, the SRB time is counted multiple times

     – Therefore *DO NOT* use the DB2 SRB time from the 101 record

Warnings

- RMF Monitor III VSAM History Data Sets
  - These data sets can contain 1110 samples (this isn't documented)
  - The size of each sample varies based on cycle time, # of active address spaces, # of devices, # of ENQs, etc.
  - RMF User's Guide recommends data sets of 50 cylinders
  - Increasing the size may not really allow you to have more history online
  - You may be only able to increase history by adding more data sets (or running less work, but let's get REAL!)
  - See **Jürgen Holtz's** session 2557 for method of calculating space

WLM Quickstart Policy

- WLM Quickstart Policy
  - This is a sample WLM goal mode policy based on my recommendations
  - I updated the article in May
  - The downloadable policy is now available, ready to modify
  - Already 556 downloads of the Quickstart Policy since May

    http://www.watsonwalker.com/quickst.html
  - This is at R3 level; if you need an earlier level, please send me an email

WLM Quickstart Policy

- NEWWORK

  - I recommended creating a service class called NEWWORK with two periods (first is short response time goal; second is velocity goal); then assign all new subsystems to that single service class and a unique report class

  - When volume of activity in NEWWORK starts to be noticeable, look at report classes to find where work is coming from; new service class may be needed

WLM Quickstart Policy

- NEWWORK
  - Unfortunately, some of the new subsystems (IWEB, LSFM, SOM) cannot handle either two periods or a velocity goal
  - New recommendation is one of the following:
    - Define NEWWORK class with single period using an easy response goal (e.g. 50% within 3 seconds) - the amount of volume should be low to begin with, so it shouldn't impact other work
    - Create NEWWORK with two periods as originally defined and also create NEWRESP (single-period using an easy response goal) for subsystems IWEB, LSFM, & SOM and any other new subsystems that can only handle one period and/or can't handle velocity goals

WLM

- **APARs**
  - Important closed APARs:
    - **OW32140** - reduces CICS V4/TS overhead if not using transaction goals; note that this does not affect users who are using transaction goals - it's still important for those users to reduce total MAXTASKS as much as possible (Chuck Hopf had a large CICS environment, 350 test regions and 40-50 production regions with a total MAXTASKS of 30,000 and found WLM taking total 40-50% of one CPU across a sysplex; reducing MAXTASKS to 8000 reduced CPU time to 16% of one CPU or 1.3% of total sysplex)
  - Open APARs at last SHARE now closed:
    - **OW37216** - higher importance server address space delayed by lower importance
    - **OW37652** - STC marked as SYST in SCHEDxx can't be assigned (R4+)
- **Keep Current With WLM Web Site**
  - **http://www.ibm.com/s390/wlm**

WLM

- **Getting to Goal Mode**
  - I have a checklist of steps to get to a monoplex, including creation of a sysplex couple data set and a WLM couple data set on my Web site (under "Articles")

  - Peter Enrico pointed out that the job step to create the WLM CDS needs to be updated to include size definitions for application environments and resources. I'll update it, but please do it yourself until then.

WLM

- ● **Goal Mode Feedback**
  - ■ Most people find it fairly straightforward and easy to implement
  - ■ Biggest confusion is for CICS and IMS
    - – How to set goals
    - – How to interpret reports
    - – How to manage test regions differently than production
    - – How to separate the SERVER address spaces from real WLM-managed work
  - ■ My latest recommendation is to let CICS and IMS run with velocity goals until you're more comfortable with goal mode

WLM

● From **Kathy Walsh's** session 2500 - OS/390 Performance "Hot" Topics:

■ A Service Definition has a 'Functionality Level'

■ Level number increases if new functions used

■ WLM Appl can process any Service Definition at a Functionality level Equal or Less than the WLM Application Level

■ All members in the Sysplex must be able to support the functionality level in use

```
   File  Utilities  Notes  Options  Help
---------------------------------------------------------------------
Functionality LEVEL001          Definition Menu          WLM Appl LEVEL008

Command ===> _____


Definition data set  . . : none

. . .
```

14

WLM

● WLM Functionality Levels

SP 5.1	LEVEL001	• Original definitions
SP 5.2	LEVEL002	• Classification rule using: collection name, correlation info, connection type, package name, plan name • Classification groups for: connection type, package name, plan name • Classification rule extensions • Service definition extensions
R3	LEVEL003	• Application environment • Classification rules using: procedure name, perform, perform group • Over 255 report classes
R4/R5	LEVEL004	• Scheduling environment • Scheduling environment resources • Classification rules using priority • Classification rules using SYSTEM or SYSSTC
R6	LEVEL006	• Description fields in classification rules or groups
R6 w/ OW33509 OR R7	LEVEL007	• Classification rules using process name
R8	LEVEL008	• Use of dynamic PAVs for ESS (Shark) • May be others…

WLM

● **WLM Functionality Levels**

■ Example of a common problem:
- Two LPARs at R4; one updated to R6; in the R6 LPAR you update the current policy by adding the new description field to the classification rules and save it; R4 can no longer read the CDS because it's now at LEVEL006 and R4 runs at LEVEL004.
- Solution: In R6, remove the description field from the policy and save it; it will save back as LEVEL004

■ Use caution when using any new function in a policy

MQSeries

- Super session handout:
  - **John B. Jones**, from IBM in the UK, session 1113, "MQSeries for MVS/ESA Performance and Tuning"
  - Download from the SHARE Web site at **http://www.share.org/proceedings/sh92/sharew99.hmtl**
  - Availability of documentation, including performance test results in the online MQSeries SupportPacs, **http://www.software.ibm.com/ts/mqseries/txppacs/**
    - MP00 - MQSeries capacity planning (includes performance test results)
    - MP15 - Program to print SMF statistics
    - MP16 - Capacity planning and tuning document
    - MP18 - MQSeries for OS/390 V2R1 - performance summary

Web Sites

- Washington Systems Center Hints & Tips
  - New section that is growing; keep an eye on this
  - TD100161 - Programs available for Viewing SMF Records
  - Lots of suggestions for IMS

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

- Amdahl CMOS List Prices
  - **http://www.eshopamdahl.com**
  - $2600/MIPS (includes 1 GB memory & 32 channels)
  - $50,000 for 512 MB of storage

Web Sites

- ● **New TASID Location**
  - ■ Doug Nadel
    - **http://www.mindspring.com/~somebody**

- ● **OS/390 User Experiences**
  - ■ R6 - Dave Alcock's Home:
    - **http://users.ticnet.com/davea/mvs/os390r6.htm**
  - ■ R7 - Francois Le Maner:
    - **http://perso.libertysurf.fr/flm**

- ● **Java Free Tutorials**
  - ■ **http://www.ibm.com/developer/java**
    - – Click Education at left for free tutorials

APARs of Interest

- **PQ28126** - TCP/IP HIPER
  - If TCPIP won't stop and a FORCE is performed, TCPIP may not restart. An IPL is necessary to clear the problem. TCP/IP V3.4 & V3.5, 99/07/19

- **OW36208** - VLF
  - Incorrectly invalidates entries causing high I/O rates. R1-R6, 99/04/08

- From **Norman Hollander** of Southern California Edison:
  - After going to OS/390 R7 and applying **OW30214**, he found that RMF post-processor wouldn't run without two new sort exits: ERBPPE15 and ERBPPE35. They're documented in the User's Guide, but no other mention was made of them and report jobs wouldn't run

- *NEW - OW40407 - Corrects WLM R6 discretionary logic*

Some Oldies But Goodies

- RMF SP 4.2 introduced new RMF Trace fields:

  - OMDGAMRE - Max msgs on AMRF queue

  - OMDGCMDI - Commands issued per second

  - OMDGOREB - Max ORES on system reply queue

  - OMDGWQEB - Max WQES on system o/p queue

  - OMDGWTOI - Lines of msgs, WTOS per second

  - OMDGWTLI - WTLs per sec (records to SYSLOG)

Some Oldies But Goodies

● **RMF Trace fields**

■ **How to use this information**

You can use the OMD fields to tune your console system. If you find a large number of messages on the AMRF queue, you could consider getting more operators to handle the volume. You can use the max ORES to determine optimum ORE buffer limits at JPL. They're changeable with the CONTROL M command. The maximum WQES can be used to determine a valid WTO buffer time limit at IPL (or use CONTROL M). The lines of messages and WTOS per second can be used for capacity purposes to determine the peak and average message rates. [Ref: CWTL, June 1991]

Some Oldies But Goodies

- **Reduce RMF Monitor Overhead**
  - You want to run RMF gatherer for data collection of sysplex and parallel sysplex, but also run another online monitor
  - How to reduce overhead of RMFGAT?
  - Increase MINTIME from 100 to 500 (this is similar to the INTERVAL parameter; records are summarized and written to the file every MINTIME)
  - This still keeps mins, maxs, and average, but you might miss a peak activity; 90% of data for sysplex is totals
  - [Ref: CWTL, August 1991]

Some Oldies But Goodies

- APARs
  - **OW32277** (overhead in SSRBs by DB2)
    - Without APAR, system overhead can be as high as 10-15%
    - **OW33027** may also help by increasing size of SSRB pool
    - [Ref: CWTL, 1998, No. 2]

- Good Reading
  - GG66-3263 - JES2 Multi-Access Spool in a Sysplex Environment
    - Has GREAT performance and tuning recommendations
    - [Ref: CWTL, Jan/Feb 1995]

HFS in V2R7

- **From Peter Enrico**
  - Session #2562 - Tuning OS/390 UNIX: A Cookbook Approach
  - Migrate to OS/390 V2R7!!
  - Significant HFS performance improvements
    - HFS deferred writing (both metadata and file changes)
      - no longer a synchronous I/O to update last reference timestamp
      - full data integrity
      - control sync interval via new SYNCDEFAULT(t) parameter

  - *Note: IBM claims up to 85% reduction in CPU time for HFS accesses*

Strongest recommendation for tuning is to migrate to OS/390 V2R7. There have been significant performance improvements to the OS/390 UNIX System Services HFS.

• HFS deferred writing

  • SYNCDEFAULT(t) - new BPXPRMxx FILESYSTYPE HFS parameter

  • specifies the number of seconds to be used as default for sync daemon interval

  • recommend 30 for important workloads

  • SYNCDEFAULT(0) turns off deferred writes

  • If SYNCDEFAULT value not 0, then exposed if system crashes. Usually OK for most types of work, but not production transaction level data.

• HFS caching in areas greater than 2BG possible

  • VIRTUAL(max) - new BPXPRMxx FILESYSTYPE HFS parameter

  • FIXED(min) - new BPXPRMxx FILESYSTYPE HFS parameter

  • System maintain I/O buffer pools

• More HFS enhancements - covered in other sessions

• New confighfs command

  • query the HFS limits in BPXPRMxx

  • change limits dynamically

  • provides additional information

HFS in V2R7

- **From Peter Enrico (cont.)**
  - HFS caching
    - pool sizes greater than 2GB possible
    - Manages caches in 4 different pools;
    - Dynamic pool expansion
    - control virtual storage HFS data and metadata buffers use via VIRTUAL(max)
    - control virtual storage fixed at HFS init time via FIXED(min)

  - Any questions: <peter.enrico@ibm.net>

HFS in V2R7

- **More from Peter**
  - **Use Temporary File System for /tmp**
  - **TFS is an in-memory physical file system that supports in-storage mountable file systems**
  - **Mount /tmp as a Temporary File System**
    - Enables /tmp to be an in-storage file system (i.e. not DASD)
    - Results in high-speed file system for temporary files
    - Could run in
      - OMVS kernel address space
      - colony address space
  - **UNIX System Services Planning Manual**
    - Chapter 12 - short chapter on TFS (temporary file systems).
    - Chapter 13.3 - setting up a colony address space.

27

Suggest mount a Temporary File system (TFS) over /tmp to improve I/O response times for temporary files.

• Add the following statements to SYS1.PARMLIB(BPXPRMxx):

```
FILESYSTYPE TYPE(TFS) ENTRYPOINT(BPXTFS)
MOUNT FILESYSTEM(/TMP) TYPE(TFS) MOUNTPOINT('/tmp') PARM('-s xxx')
```

where 'xxx' is the number of megabytes you want to allocate to the in storage file system mounted over /tmp

Notes:

• Maximum file size of 2GB

• When system IPLed or crash - lose all the /tmp files. (Should be OK with users)

• TFS normally runs in kernel address space (OMVS)

• Optionally set up a 'colony' address space for TFS to run in

```
FILESYSTYPE TYPE(TFS) ENTRYPOINT(BPXTFS) ASNAME(TFSCOLON)
MOUNT FILESYSTEM(/TMP) TYPE(TFS) MOUNTPOINT('/tmp') PARM('-s xxx')
```

Where TFSCOLON is a procedure you set up in SYS1.PARMLIB to EXEC PGM=BPXVCLNY

Questions

- **Email: cheryl@watsonwalker.com**

- **Web site: http://www.watsonwalker.com**

- **CD-ROM: http://www.watsonwalker.com/TUNING.html**