

Cheryl's Hot Flashes #11

Cheryl Watson

Session 2509; SHARE 102 in Long Beach, CA

February 27, 2004

Watson & Walker, Inc.

home of "Cheryl Watson's TUNING Letter", BoxScore and GoalTender

Ain't it the Truth!



- Found on IBM-Main:

"Give a person a fish and they will eat for a day, teach them to use the Internet and they'll leave you alone forever..."

- Taglines
 - "Being a computer means never having to say you're sorry." (Jerry Urbaniak)

Agenda



- Survey Questions
- Interesting Items
 - SMF Type 30 CPU time, Neat Info, APARs
- User Experiences
 - 31-bit on zSeries, zSeries, ECMBs
- 6-Month Update
 - On/Off Capacity on Demand, Trace, Spin Loop
- This SHARE
 - New LSPR workloads, ICKDSF

Survey Questions



- Who is planning to investigate On/Off Capacity on Demand in the next year?
- Who is running z/OS.e?
- Who is using Variable WLC pricing?
- Who has activated IRD CPU management?
- What is your next upgrade:
 - z/OS 1.4
 - z/OS 1.5
 - z/OS 1.6
 - Other?

Survey Questions



- Is bimodal testing important?
 - z/OS 1.4 is last release to support bimodal
 - Please categorize yourselves under just one of the following:
 - Tested or will test 64-bit under OS/390 R10
 - Tested or will test 64-bit under z/OS with bimodal support
 - Went or will go to 64-bit under z/OS directly
 - without bimodal support at all
 - with bimodal support as a backup
 - Will not go to 64-bit
 - For those who tried bimodal support in either z/OS or OS/390, did you need to fall back because of problems?

Interesting Items - SMF



- SMF CPU Time
 - Lack of SMF precision in SMF type 30 records
 - Discrepancy between service units and CPU time
 - Recent APAR
 - New SMF CPU field

Interesting Items - SMF



- Lack of SMF precision in type 30
 - SMF30CPT (TCB), SMF30CPS (SRB)
 - Data is in hundredths of seconds (hh:mm:ss.hh)
 - Hundreds of job steps on z900 had CPU times of .01 second. On z990 they had either .01 (so charge back was the same) or 0 (and charge back dropped)
 - Primary billing basis for most installations

Interesting Items – SMF



- So what about service units?
 - SMF30CSU (TCB), SMF30SRB
 - Data is in service units (CPU time multiplied by service definition coefficient multiplied by service units per second for the machine)
 - Smallest service unit value on a z900 is about 50 (or .0005 seconds)
 - Problem – while the service units per second value is in type 30 record (RMCTADJC), the service definition coefficient is not, so you can't calculate CPU time from service units unless you also process the type 72 records

Interesting Items - SMF



– z900 example of CPU times (hh:mm:ss):

Description	Number of Records	% of Records	Total CPU Time	Difference: SU - CPU
Total TYPE30_4	1,069,999	100.0%		
CPU = 0, SU = 0	216,815	20.3%	0	
CPU = 0, SU > 0	399	0%	00:00:02	00:00:02
CPU > 0, SU = 0	8,671	.8%	00:08:48	-00:08:48
CPU <= .1 sec	556,800	52.0%	06:14:40	00:01:28
CPU <= 1 sec	218,552	20.4%	18:12:45	00:01:35
CPU <= 1 min	60,681	5.7%	138:39:05	00:17:55
CPU > 1 min	8,065	.8%	1077:45:24	-04:13:24
Largest 16 steps	16	0%	404:27:53	-285:27:07

Interesting Items – SMF



- Problems Due to Lack of Precision
 - In our example, less than 1% of the jobs account for the majority of the CPU time; but 4.25 hours of chargeable time (\$10,000 or so) is lost due to lack of precision in SMF CPU times
 - In our example, 52% of job steps will become zero or vary by 10% to 100% when moving to a z990
 - The data is not available in the SMF type 30 records to calculate CPU time from service units
 - Charge back will be inconsistent after an upgrade
 - Comparison of CPU time inconsistent after an upgrade

Interesting Items - SMF



- Solution to Lack of Precision
 - Change definition of CPU fields to have more precision (unacceptable)
 - Change definition of SU fields to be raw service units so SDC is not required (unacceptable)
 - Add new CPU time fields in microseconds (just TCB and SRB or all?)
 - Add SDC so CPU time can be calculated from service units
 - Ignore it; it's too small to bother with?

Interesting Items – SMF



- APAR [OA06407](#) – Large or negative values in SMF30SRV
 - Closed 18Feb2004 for OS/390 R10+
 - **Kathy Walsh** mentioned this in session 2500
 - What I saw: same z900 example as before
 - The 16 largest job steps out of 1,069,999 had CPU time greater than the service unit time (404 hours CPU time versus 119 hours SU time)
 - So before using service units, be sure to apply the fix for this APAR!

Interesting Items – SMF



- z/OS 1.3 added new field – SMF30CEP
 - CPU time consumed for an address space or job while enqueue promoted
 - But it's an accumulative field as the job runs, so subsequent step records (subtype 4) carry over the values from previous steps; ditto for interval records
 - IBM rejected an APAR to correct it saying that it's working as designed
 - We think it was designed wrong
 - Thanks to **Darrel Faulkner**, director of the NeuMICS Development team at CA for pointing this out

Interesting Items – Neat Info



- APAR [OW41269](#)
 - Documentation For System Managed Buffering (SMB), 09Nov1999
 - Excellent guidelines and restrictions for SMB
 - Was supposed to be added to Using Data Sets manual, but hasn't made it yet
 - Thanks to **Michael Friske** for pointing this out

Interesting Items – Neat Info



- **Michael Raithel's** - second edition of "Tuning SAS Applications in the OS/390 and z/OS Environments"
 - Available from Amazon or SAS publications for about \$50 -
www.sas.com/apps/pubscat/welcome.jsp
 - Savings from its first use can justify the cost of the book

Interesting Items – Neat Info



- Great Redbook!
 - **SG24-5952** - z/OS Intelligent Resource Director
 - Great section on how LPAR dispatching works
- Hex, Octal, Binary Calculator
 - **www.google.com**
 - Enter mathematic calculation into google search
 - Prefix hex numbers with 0x, octal with 0o, binary with 0b
 - E.g. 0x104F + 0x2072
 - Google returns with 0x104F+0x2072=0x30C1

Interesting Items – Neat Info



- WSC Papers
 - **PRS752** - Performance Summary Report for SMF 120 records from WAS V.5 for z/OS - this is a free Java tool that produces a summary report showing activity for each J2EE server instance, bean and method from the SMF 120 records (03Nov2003)
 - **TD101238** - zSeries Performance: Determining the Logical CP Requirements for a Partition - like any of **Kathy Walsh's** papers, this one is a MUST-READ (03Oct2003)

Interesting Items – Neat Info



- Console Performance Paper
 - Console Performance Hints and Tips for a Parallel Sysplex Environment
 - www.ibm.com/zseries/library/techpapers/pdf/gm130166.pdf
 - Excellent section on why Console address space can have high CPU spikes
 - Thanks to **Paul Zielstra** from Acxiom

- [OA03870](#)
 - OS/390+, 24Sep2003
 - Inits At The Enqueue Promotion Priority When INITIMP Specified
 - When specifying a value for INITIMP in the IEAOPTxx parmlib member, the actual dispatching priority of the idle initiators may be higher than the CPU critical work at the importance specified in the INITIMP keyword.

- [OA04315](#)
 - z/OS 1.2+, 26Nov2003
 - Unnecessary Paging While Dumping (SVC Dump) D/T2064 and SVT Corrections
 - Performance degradation due to SDUMP processing. Possible performance degradation due to application frames being stolen before frames backing SDUMP captured pages, which are better steal candidates. Flagged as HIPER PERFORMANCE. Requires [OA04323](#).

- [OA05216](#)
 - OS/390 R10-z/OS 1.5, 19Nov2003
 - Temporary Data Sets not Deleted After **OA02195**
 - [OA02195](#) was taken to address high CPU utilization when a long-running job dynamically allocates many temporary data sets. After the application of the PTFs for **OA02195**, a temporary data set created in one step of a job as DISP=(NEW,PASS) and referenced in a later step of the job as DISP=(OLD,DELETE) does not get deleted. It doesn't matter whether the data set has a system generated name (DSN=&&dsname or DSN omitted) or a permanent name. So STORAGE volumes may fill up with data sets which should have been deleted. Flagged as **HIPER FUNCTIONLOSS PERVASIVE**.

User Experiences – 31-bit on zSeries



- 31-Bit on zSeries
 - Common scenario
 - OS/390 R10 running on non-zSeries and upgraded to zSeries
 - Normally storage increase is very large
 - Example
 - Site moved from 9672 with 1 GB of storage to z800 with 8 GBs of storage (minimum configuration)
 - CPU time in *MASTER* increased from 1% to 3% (we've seen this higher in other sites)
 - Cause is due to additional scanning of expanded and frequency of central scanning (every 2 seconds)

User Experiences – 31-bit on zSeries



- 31-Bit on zSeries
 - Solution
 - Vary storage offline until it's needed
 - Migrate to 64-bit using bimodal mode
 - 64-bit will eliminate scanning of expanded and frequency of central storage scanning decreases to every 10 seconds
 - Use SoftCap to estimate cost of move

User Experiences – zSeries



- Speaking of Larger Storage
 - RSU=xx
 - Parameter in IEASYSxx
 - Used to be in storage increments
 - In OS/390 R10, you can also specify MB, GB, %
 - Please, please, please stop using increments
 - RSU=2 means:
 - 32 MB on smallest z800
 - 128 MB on z900 with 32 GB
 - 256 MB on z990 with 40 GB

User Experiences – ECMB



- z/OS 1.4
 - Exploitation Support Feature
 - Replaces CMB with ECMB
 - ISVs may needed updates
 - Until then, you can turn off with ECMB=NO in IEAOPTxx
 - See WSC Flash10270, 09Jan2004

6-Month Update - OOCoD



- Daily On/Off Capacity On Demand
 - On z990, you can add a CPU dynamically and only pay for the number of days you use it
 - IBM reduced the On/Off CoD Software Charges for z990 from 30% and 50%
 - Announcement 304-004 on 13Jan2004

6-Month Update - Trace



- DFSMS Trace
 - We suggested turning this off in our Nov/Dec 1993 TUNING Letter
 - Default is TRACE(ON), which is IBM's recommendation
 - But CPU overhead is incurred (one site saved about 45 minutes of CPU time a day by setting trace off)
 - To see if it's on, use operator command: `d sms,trace`

6-Month Update – Spin Loop



- Spin Loop
 - A site experienced several spin loops (and IPLs) because fixed frames filled up storage between 16 MB and 2 GB; SRM did not detect this critical storage shortage
 - This condition can occur in any site with lots of real storage and large enough CPs to keep the system very busy
 - APAR [OA03577](#) - Spin Loop Abend071 IAXUA Fixed Storage Between 16Meg and 2Gig, 7/30/03; changed to HIPER after it came out (OS/390 R10-z/OS 1.4)

This SHARE



- Favorite Session - 2513 – Introduction to Processor Sizing, **Joanne Brown**, WSC
 - New LSPR workloads for zSeries:
 - CB-L – Commercial Batch (similar to CBW2)
 - CB-S – Commercial Batch Short (similar to CB84)
 - OLTP-W – Online (similar to CICS/DB2)
 - OLTP-T – Online (similar to IMS)
 - WASDB – WebSphere

- Session 2513 - Processor Sizing (cont.)
 - Marketing reps can use internal tools for processor sizing studies (zPCR and CP2000)
 - Three new custom workloads:
 - Online: $(\text{OLTP-W} + \text{OLTP-T} + \text{WASDB}) / 3$
 - Other: $(\text{CB-L} + \text{CB-S}) / 2$
 - Low I/O: $(.60 * \text{CB-L}) + (.20 * \text{OLTP-W}) + (.20 * \text{WASDB})$
 - Low I/O is defined as images that have fewer than 30 DASD SSCHs/Second per MSU

This SHARE



- **Harv Emery**, Session 2867, LPAR Advanced Topics
 - ICKDSF 17
 - Required for all images that share DASD with a z990
 - This ICKDSF level must be installed on all sharing images before z990 is installed
 - Instructions/Macro Changes
 - Store Channel Path Status (STCPS) is now a no-op on z990
 - Store CPU ID STIDP has new format on z990 (no logical CPU id) – use Store CPU Address (STAP)

Questions



- **Email:**
 - cheryl@watsonwalker.com
- **Web site:** www.watsonwalker.com