

Adjusting RMPTTOM to Reduce SRM Overhead

Kevin Martin – McKesson

Last December Kevin Martin modified RMPTTOM to invoke SRM (System Resource Manager) less frequently and reduce system overhead.

Kevin will summarize some discussions from the MXG listserver about this topic, explain APAR OA18452 which adjusts the RMPTTOM default for some fast processors, and explain how he quantified the reduction in CPU even though SRM activity is not captured.

This is intended to be an introductory level presentation about a component in the z/OS operating system.

What is RMPTTOM?

- **RMPTTOM** Specifies the SRM invocation interval. The specified real-time interval is adjusted by relative processor speed to become SRM time in order to ensure consistent SRM control across various processors. The relationship of real time to SRM time for each processor is described in the “Advanced SRM Parameter Concepts” section of *z/OS MVS Initialization and Tuning Guide*.
- Some of the functions within SRM (and WLM) are a function of elapsed time (e.g., the WLM policy adjustment interval currently takes place every 10.24 seconds). Some memory controls are based on multiples of the SRM second.

MXG listserver discussion

- Probably the most exciting revelation I have encountered work-wise this century.

Geoff Adams Sep 27, 2006

- Not a new idea: the default SRM invocation interval was changed from 500 to 1000 milliseconds (July 1985)
Don Deese
- All z/OS users owe Geoff Adams at National Australia Bank three cheers.
Bernie Pierce Oct. 24, 2006

Go to www.mxg.com and search listserver archives for RMPPTOM

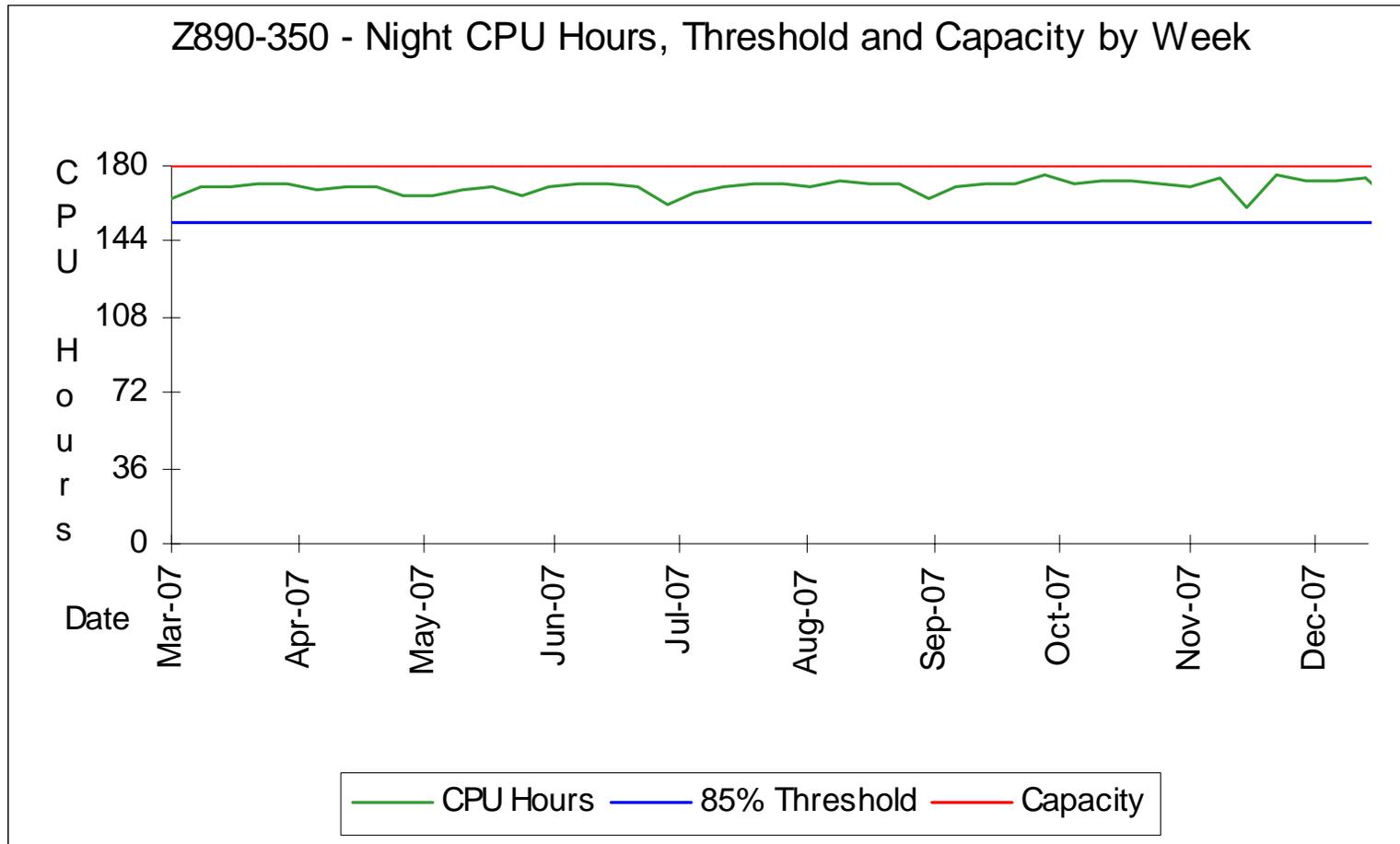
Bernie Pierce recommendations

- By the time the z990 was introduced the SRM second was reduced to about 2 milliseconds or nearly 500 timer interrupts per second.
- Multiple LPARs is very key to the opportunity to reduce SRM CPU cost. A major factor determining SRM CPU cost is the number of address spaces swapped in.
- Some LPARs have little need for SRM management: swap analysis, period switch, mean time to wait etc.
- set RMPTTOM=2000 on a z9 EC machine. Also a reasonable minimum for a z990 or z900 as well. For non-production LPARs higher values may be acceptable.

APAR OA18452

- IBM provided PTFs to address rounding that doubled the SRM timer rate on z9 vs z990.
- **Default RMPPTOM Value:** 3000 (for systems with a uni-processor speed of more than 100 MIPS) 1000 (for systems with a uni-processor speed of 100 MIPS or less)
- The installation's ability to save CPU time will depend upon different configuration factors. The amount of uncaptured time is a function of:
 - the SRM interval (influenced by RMPTTOM)
 - number of Address Spaces in the LPAR
 - number of LPARs

Motivation: Busy night shift



Resetting RMPPTOM values

- Using 2086 model 350 and model 250, approximately 150 MIPS per CPU engine
- Tried to use conservative values. Default was 1000.
For Production, set RMPPTOM=3000
For Test and QA, set RMPPTOM=6000
- Activated dynamically: SET OPT=xx
- Used monitors to measure increase in production CPU utilization while test LPARs were idle.
- The increase in production activity = savings from test. Production CPU utilization increased from 89% to 92%.

Display Filter View Print Options Help

```

SDSF DA KZ01 DDCA PAG 0 CPU/L 100/ 93 LINE 1-19 (275)
COMMAND INPUT ==> _ SCROLL ==> CSR
NP JOBNAME ProcStep C Pos DP SrvClass SP CPU-Time CPU% SIO EXCP-Cnt
DY65J2CE DY65G05 5 IN DD PRDBATA3 2 265.61 20.20 399.71 182,768
DY65J2CF DY65G05 5 IN DD PRDBATA3 2 452.11 20.20 409.71 266,386
DY65J2CD DY65G05 5 IN DD PRDBATA3 2 140.35 19.24 489.65 110,734
DY65J2CA DY65G05 5 IN DD PRDBATA3 2 402.06 18.92 319.77 238,605
COPYCROS COPYCROS NS F5 STCMED 1 434737.88 6.09 579.58 704M
EIUJK1C N124 IN F7 TSDPRD 1 764.83 3.21 0.00 736,614
OL04J1ED OL04G05 H IN DD PRDBATA3 2 24.52 2.24 609.56 238,437
OL04J1EB OL04G05 H IN E1 PRDBATA3 1 0.81 0.96 299.78 5,842
MIMTAPE MIMGR NS FE SYSSTC 1 3482.80 0.64 0.00 1,819
RMOSTC RMOSTC NS FE SYSSTC 1 15665.81 0.32 0.00 72M
TCP/IP TCP/IP NS FE SYSSTC 1 48739.08 0.32 0.00 4,529
MQ8TMSTR MQ8TMSTR NS FE SYSSTC 1 1203.38 0.32 0.00 51,937
DB2ODIST IEFPROC NS FB STCHI 1 13358.33 0.32 0.00 134,364
*MASTER* NS FF SYSTEM 1 89037.05 0.00 0.00 3,191,696
PCAUTH NS FF SYSTEM 1 0.09 0.00 0.00 16
RASP NS FF SYSTEM 1 1063.42 0.00 0.00 2
TRACE NS FF SYSTEM 1 0.14 0.00 0.00 70
DUMPSRU DUMPSRU NS FF SYSTEM 1 52.44 0.00 0.00 370,850
XCFAS IEFPROC NS FF SYSTEM 1 16940.52 0.00 0.00 17M

```

JOBNAME: TMONMUS LOGICAL PARTITION SUMMARY DATE: 04/12/08

SYSID : KZ01 TIME: 10:39:18

COMMAND: _____ CYCLE: MMSS

SRM STATS: CPU= 99.81 UIC=2540 TPR= 0 SIO=1087.1

START: 04/12/2008 10:39:00 DURATION: 00:18.44 TYPE: DELTA VIEW: 1

LPAR CLUSTER NAME: PRDPLEX WEIGHT MGMT: DISABLED

CURRENT PART: DDC1LP01 PROC RUN TIME: DYN WAIT COMP: N *PHYS: Y

CPU		<--- DISPATCH UTIL (%) -->	<----- LPAR MGMT (%) ----->
CATEGORY	#	0---20---40---60---80--100	0---20---40---60---80---100
SHARED	3	99.20	0.50
DEDICATED	0	0.00	0.00
ALL CPS	3	99.20	0.50
ICF/IFL/IFA	1	99.70	0.00

<--- LOGICAL PARTITION ---->							<----- DISPATCH UTILIZATION (LPAR) ----->			
NAME	STA	#LP	CAP	CUR	TGT		<---- LOGICAL (%) ---->		<--- PHYSICAL (%) --->	
	A/I	D/S	STA	WGT	SHR		0--20--40--60--80--100	0--20--40--60--80--100		
DDC1LP01	A	0 3	NO	270	51		91.50	91.50		
DDCPRD1	A	1 0	NO	ICF	N/A		99.70	99.70		
DDCQAD1	A	0 1	NO	89	17		1.50	0.50		
DDCTST1	A	0 1	NO	89	17		1.60	0.50		
DDC1LP02	A	0 1	NO	52	10		6.20	2.00		
DDC1LP03	A	0 2	NO	14	3		3.90	2.60		

HELP INFORMATION = PF1

TMU08

PF KEY ASSIGNMENTS = PA1

Conclusions

- Saved approximately 3% of overall CPU utilization in our environment by using conservative RMPPTOM settings to reduce SRM overhead.
- No obvious impact to response time.
- Worthwhile if you have a constrained system