Top New z/OS Performance Functions Every Sysprog Should Understand

Glenn Anderson IBM Systems Lab Services



#### May 2018



Topics to be Covered:

- Pervasive Encryption and zBNA
- RMF Updates
- WLM Updates
- SMF Updates



#### Washington Systems Center Techdocs

www.ibm.com/support/techdocs



TRM

3

# Pervasive Encryption and zBNA



## Data Protection / z/OS Dataset Encryption





#### Data set encryption

#### • DFSMS:

- -Provides enhancements to SMF 42.6 records
- Used by zBNA to perform capacity planning estimation on data set encryption costs

z/OS release Function	V2.1	V2.2	V2.3
DFSMS SMF 42-6 Enhancements	OA52132	OA52132	OA52734

Release	z9 EC / z9 BC	z10 EC / z10 BC	z196 / z114	z13 / z13s	z14
z/OS 2.1	YES – PE NO	YES – PE NO	YES	YES	YES
z/OS 2.2	NO	YES – PE NO	YES	YES	YES
z/OS 2.3	NO	NO	NO	YES	YES

### Data protection / Coupling Facility encryption



## Coupling Facility structure encryption



- Provide new measurements to capture additional information on the data read/written to list/cache structures
- -Used by zBNA to perform projections on data set encryption costs

RMF

- Provide new fields in SMF 74.4 records to capture new structure information provided by XES in APAR OA51879
- -Updates to the RMF CF Post-processor report
- -Used by zBNA to perform projections on data set encryption costs

z/OS release Function	V2.2	V2.3
XES Support for Enhanced List/Cache Reporting	OA51879	Base
RMF SPE for enhanced List/Cache Reporting	OA52003	Base

## IBM Z Batch Network Analyzer

- IBM z Systems Batch Network Analyzer

   A no charge, "as is" tool to analyze batch window resources, and QSAM/BSAM zEDC Compression candidates
   Available to Customers, Business Partners, and IBMers

  - PC based, and provides graphical and text reports
     Including Gantt charts and support for Alternate
    - Processors
- Updated to provide information on Pervasive Encryption -zBNA V1.8.1
  - -Available on 8/25/17
- Available on Techdocs
  - Customers
    - https://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5132
  - Business Partners
    - https://www-
      - 356.ibm.com/partnerworld/wps/servlet/mem/ContentHandler/tech PRS5133
  - IBMers
    - http://w3-03.jbm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5126



									n	ming								
	00.15	00:20	00.45	01:00	01:15	01:30	01)45	02:00	02:15	02:30	02.45	03.00	03:15	03:30	03.45	0400	04:15	04:30
DEVENDER	-																	
AT DEVELABLE	Sec.		-															
0538030																		
ALDSVEDDY		_																
05YHL31F																		
AT DOWALDIE		10.00																
05197347		-																
AT DEV#F34F		100																
03Y#122F			-															
A1 DSV#I22F																		
DSYRODUO		- 1																
A1 DSY86386																		
DSYMHOSK					_													
AT DEVENDER																		
DSYRIDOL																		
AT OSYMIDOL																		
DSYMHOUR																		
AT DEVENDER				10.00		- 12												
DEVIFICO						122												
AT DSYMPTOG				100														
DSWECTOR																		
A1 DSYNCPOD				10.00														
DSY#6386																		
A1 D57#0300				100														
ND8#P100																		
A1 NOB#P100					100	100	1.1	2.2	1.1	1			2.2	11	1.1	100	1.1	1212
DEVALUE																		
AT DEVALUEL					11													
05Y#1226																		
A1051#1220					12.0													
DSYNCPOB																		
A1 DEVICEOR					10.00	11												
DSYMHOOD																		
A1 DSYMHOUR					_	1.1	12.	11	11	1.								

#### zBNA – multiple data sources

IBM







Integrating multiple data sources to provide cutting edge analytics for batch workloads

#### zBNA – a PC based tool



#### IBM

Multiple Factors which influence Encryption Costs

- –Server technology
- -Number of bytes involved in the operations
- -Number of calls made to the encryption service
- -Read:Write ratio of the data

 Compression technologies reduces number of bytes in the data sets and indirectly influences encryption costs

#### zBNA encryption home panel



#### zBNA estimated data set encryption CPU time



## **RMF Updates**



#### **RMF Product Overview**



#### IBM

**RMF** enhancements for CF Encryption

- RMF V2.2 and V2.3 stores new metrics on Coupling Facility read and write statistics for every CF structure in SMF type 74.4 record
- RMF V2.3 enhances Postprocessor CF Activity report and Monitor III CFACT report to display the encryption state of each CF structure.

z /	/OS V2R3	SYSPLEX RPT VERS	C O U SYSDP ION V	PLI PLEX 2R3 RMF	NG	FACIL DATE 02 TIME 09	I T Y /03/2017 .59.35	ACTI	VITY INTERV CYCL	AL 015.00 01.000 SE	CONDS	р	PAGE 1
COUPLI	NG FACILITY NAME SAMPLES (AVG) -	- CF02 900 (MAX) -	900	(MIN)	- 9	900							
				COUP	LING	FACILITY	USAGE	SUMMARY					
GENERA	AL STRUCTURE SUMM	ARY											
TYPE	STRUCTURE NAME	STATUS CHG	ENC	ALLOC SIZE	% OF CF STOR	* REQ	≉ OF ALL REQ	% OF CF UTIL	AVG REQ/ SEC	LST/DIR ENTRIES TOT/CUR	DATA ELEMENTS TOT/CUR	LOCK ENTRIES TOT/CUR	DIR REC/ DIR REC XI'S
LIST	IXCPLEX_PATH3	ACTIVE	YES	12M	0.8	2202	7.5	4.8	2.45	987	950	N/A	N/A
	IXCPLEX_PATH4	ACTIVE	NO	12M	0.8	25352	86.1	17.2	28.17	987	950	N/A	N/A
	IXCVLF	ACTIVE	YES	12M	0.8	1876	6.4	1.7	2.08	987 1	950 16	N/A N/A	N/A N/A
LOCK	IGWLOCK00	ACTIVE SEC	N/A	7 9M	5.1	0	0.0	66.8	0.00	184K 0	0	8389K 0	N/A N/A
	STRUCTURE T	OTALS		115M	7.4	29430	100	90.5	32.70				

\* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

ONSDOOS

#### z14 - Support of Crypto Express6S Card

	c	z/0S	V2R2	ODBOCES	SYSTEM ID S35 RPT VERSION V	2R2 RMF	DATE 04/26, TIME 09.00	/2017 .00	INTERV CYCLE	AL 29.59.848 1.000 SECONDS	
	C	KYPI OGR	APHIC CCA C	OPROCES							
CEX5C	ID 0 1 5 6 7	RATE 1882 1927 3479 3235 3504	EXEC TIME 0.206 0.201 0.111 0.120 0.110	UTIL% 38.8 38.8 38.6 38.7 38.7	RATE 0.25 0.24 0.46 0.45 0.43				C re of C	rypto Express6 Ca ported with same measurements a rypto Express5	ard set as
·	CF	RYPTOGR	APHIC PKCS1	1 COPRO	CESSOR					;	
			TOTAL			OPERATIONS D	DETAILS		/		
YPE	ID 2	RATE	EXEC TIME	UTIL%	FUNCTION	RATE	EXEC TIME	UTIL%			
CEXSP	2	549.0	0.011	55.0	ASYM FAST	209.4	0.373	10.0	V		
					ASTM GEN	219 3	0 651	14 3			
					SYMM COMPLET	F 4.45	0.542	0.2			
					SYMM PARTIAL	36.27	0.663	2.4			
CEX6P	8	659.2	0.508	33.5	ASYM FAST	331.3	0.464	15.4			
					ASYM GEN	0.13	1.203	0.0			
					ASYM SLOW	260.0	0.552	14.4			
					SYMM COMPLET	E 5.39	0.291	0.2			
					SYMM PARTIAL	. 62.41	0.571	3.6			
	CF	RYPTOGR	APHIC ACCEL	ERATOR							
			TOTAL		ME-	FORMAT RSA OF	PERATIONS	CRT-F	ORMAT RSA OPE	RATIONS	
ТҮРЕ	ID	RATE	EXEC TIME	UTIL%	KEY RATE	EXEC TIME	E UTIL%	RATE	EXEC TIME	UTIL%	
CEX6A	4	13504	0.029	38.7	1024 4361	0.010	4.4	1439	0.034	4.9	
					2048 1633	0.026	4.2	0.00	0.000	0.0	
					4096 6072	0 042	25 2	0 00	0 000	0 0	

#### IBM zHyperLink Express - Overview

#### What is IBM zHyperLink™?

- zHyperLink Express is a direct connect short distance IBM Z I/O feature designed to work in conjunction with a FICON or High Performance FICON SAN infrastructure
- IBM zHyperLink<sup>™</sup> dramatically reduces latency by interconnecting the z14 CPC directly to the I/O Bay of the DS8880
- zHyperLink is FAST enough the CPU can just wait for the data
  - No Un-dispatch of the running task
  - No CPU Queueing Delays to resume it
  - No host CPU cache disruption
  - Very small I/O service time
- Operating System and Middleware (e.g. DB2) are changed to keep running over an I/O



#### **RMF Postprocessor PCIE Activity report**



#### RMF Postprocessor PCIE Activity report ...

Connectivity data and
request performance
measurements on LPAR
and CPC level

#### Synchronous I/O Link Activity

Function ID	Function CHID	Port ID	Serial Number	Type-Model	Total Request	Total Request	Successful Request %	Successful Request %	Read Transfer	Read Transfer	Read Transfer	Read Transfer Datio	Write Transfer Data	Write Transfer Data	Write Transfer Datio	Write Transfer Datia	Time Busy	Time Busy
Ļ	t	lt 1	t L1	· ↓	Rate ↑ ↓	rate † (CPC) ↓†	↓†	(CPC) ↓1	Rate ↓1	(CPC) ↓↑	ratio ↓1	(CPC) ↓↑	rate ↓1	rate (CPC) ↓	Ratio ↓1	(CPC) ↓↑	7₀ ↓'	↑ (CPC) <u>↓</u> ↑
0300	01C8	1	000000YT111	002107-981	791	6277	100	100	3.24	25.7	0.004	0.004	0	0	0	0	1.41	11.4
0301	01C8	1	000000YT111	002107-981	791	6277	100	100	3.24	25.7	0.004	0.004	0	0	0	0	1.41	11.4
0302	01C8	1	000000YT111	002107-981	791	6277	100	100	3.24	25.7	0.004	0.004	0	0	0	0	1.41	11.4
0303	01C8	1	000000YT111	002107-981	791	6277	100	100	3.24	25.7	0.004	0.004	0	0	0	0	1.41	11.4
0304	01C8	2	000000YT111	002107-981	791	6277	100	100	3.24	25.7	0.004	0.004	0	0	0	0	1.57	12.6

#### Synchronous I/O Response Time Distribution

	Function ID	% Read <	% Read <	% Read <	% Read <	% Read <	% Read <	% Read <	% Read <	% Read <	% Read >=100usec
I/O Response Time	↓†	20usec <u>↓</u> †	30usec $\downarrow\uparrow$	40usec ↓†	50usec ↓†	60usec ↓†	70usec <u>↓</u> †	80usec <u>↓</u> †	90usec ↓†	100usec ↓†	.↓†
Distribution for	0300	82.5	17.5	0.026	0.003	0.004	0.003	0.002	<.001	0	<.001
instructions executed	0301	82.4	17.5	0.026	0.009	0.004	0.004	0.003	<.001	0	<.001
synchronously on	0302	82.4	17.6	0.031	0.006	0.006	0.005	0.003	<.001	<.001	<.001
zHyperLinks.	0303	82.3	17.7	0.023	0.007	0.003	0.003	0.001	<.001	0	0.001
	0304	22.3	77.6	0.082	0.008	0.004	0.004	0.003	<.001	0	0.001

#### **RMF Postprocessor Device Activity report**



#### **RMF Monitor III PCIE Activity report**

0.004

0

0

 Monitor III PCIE Activity report enhanced to report IBM zHyperLink activities

									RMF Syncl	hronous I/(	O Link Activity	
	RMF	v2r3	PCIE A	ctivity				Line	Function ID : Allocated :	0500 22.56.47	Alloc Time % on 05/03/17	: 100
Samples: 60	System:	та5	Date:	05/04/1	7 ті	me: 02.0	06.00	Range:	Synchronous I/	0 Link		
ID CHID Type	- Function Jobname	ASID	Status	Alloc Time%	- PCI Load	Operati Store E	ons R Block	ate Refr	Port ID Type-Model Serial Number	: 1 : 002107 : 000000	-981 0YT111	
0078 01FC ZEDC 007F 01FC ZEDC	FPGHWAM	0013	Alloc	100 100	0 0	0 0	0 0	0 0	Adapter	Thi	s Function	Link (CEC)
0500 01C8 2Hyp 0501 01C8 2Hyp 0502 01C8 2Hyp	L IOSAS L IOSAS	001A 001A 001A	Alloc Alloc	100 100 100	H	IBM zH	yperLi	nk	Time Busy % :	1	. 38	11.0
0503 01С8 zнур	LIOSAS	001A	Alloc	100					Request Rate : Success % :		746 100	6022 100
Cursor sensitiv	ve fields F wing pop	<sup>-</sup> uncti o-ups:	ion ID,	CHID a	and <sup>-</sup>	Type di	splay	/	Transfer Read Rate :	3.0	06	24.7

- RMF Hardware Accelerator Activity
- RMF Hardware Accelerator and Compression Activity
- RMF RoCE Activity
- RMF Internal Shared Memory Activity
- RMF Synchronous I/O Link Activity

\* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Read Ratio :

Write Rate :

Write Ratio :

0.004

0

0

#### WLM transaction level reporting to Identify Mobile/Cloud Workloads

MWP and zWPC offer a discount on MSUs consumed by transactions that originated from a mobile device or new public cloud workloads

To take advantage of this discount, you need a process, agreed upon by you and IBM, to **identify (tag and track)** mobile- or cloud-sourced transactions and report on their consumption

- → NEW Identify mobile or cloud transaction via a transaction level attribute in the WLM service definition
  - Processor consumption data aggregated by WLM
  - · Reporting integrated into standard performance monitors (RMF) and low volume SMF records
  - Applicable to wide range of workloads, including enclave work and CICS/IMS work



#### WLM transaction level reporting to Identify Mobile/Cloud Workloads

- In your WLM classification rules, classify transactions as MOBILE, CATEGORYA or CATEGORYB
- The assigned attribute is independent from the assigned service and report class
  - Eliminates the need for using new dedicated classes for mobile or cloud workload reporting
- The assigned attribute is transparent to subsystems
- WLM tracks and reports the total and the MOBILE, CATEGORYA and CATEGORYB CPU consumption for all service and report classes
  - With exploiting levels of CICS and IMS, CPU consumption data is also available for CICS and IMS transaction service and report classes that previously did not report any CPU consumption data
  - Subsystems using independent enclaves can participate transparently; only the classification rules need to be updated.
- WLM also aggregates and reports the system-wide MOBILE, CATEGORYA and CATEGORYB consumption data

#### **RMF sample Workload Activity Report**



Get processor consumption data for CICS and IMS transactions from **low-volume** SMF records:

PORT BY:	POLICY=	BASE	WORKLOAD=CICSWKLD	SERVICE CLASS= CRITICAL DESCRIPTION	CICSFAST =NONE =CICS	RESOUN	RCE GROUP=*NONI
-TRANSA	CTIONS-	TRANS-TIME	E HHH.MM.SS.TTT				
AVG	0.00	ACTUAL	1.432				
MPL	0.00	EXECUTION	710				
ENDED	632	QUEUED	0				
END/S	10.58	R/S AFFIN	0				
#SWAPS	0	INELIGIBLE	Ξ Ο				
EXCTD	631	CONVERSION	1 0				
AVG ENC	0.00	STD DEV	0				
REM ENC	0.00						
MS ENC	0.00						
TRANSAC	TIONS AP	PL% : TOT MOE	TAL : CP 185.92 BILE : CP 99.16	AAP/IIP ON CP AAP/IIP ON CP	0.00	AAP/IIP AAP/IIP	0.00 0.00

#### **RMF sample Workload Activity Report**

Check **region overhead** of CICS and IMS regions:



# WLM Updates



## WLM & SRM enhancements for Spark

- IBM is enabling the Apache Spark platform on z/OS. The z/OS Platform for Apache Spark ٠ offers an enterprise-grade solution to enable Apache Spark natively on z/OS.
- Spark is an in-memory compute engine for analytics, operating on data in memory. ٠
- Problem: •
  - Since Spark can be very processor and memory intensive customers need a means to limit Spark's consumption of both resources.
  - Spark is primarily running on specialty engines but WLM's resource groups cannot be used to limit processor access to zAAPs and zIIPs.
- Solution: ٠
  - WLM provides new service class attribute that allows to prohibit specialty-engines eligible work to overflow to regular CPs for needs help processing. Honor Priority = DEFAULTINO
    - Limit Sparks' processor consumption to speciality engines
  - WLM provides new data field for resource groups that allows to specify an upper limit for real memory consumption.
    - Limit Spark's memory consumption

**RSM APAR OA51171** 







Memory Limit = xxx GB

#### IBM

#### Honor Priority by Service Class

Command ===>	Modify a Service Class	Row 1 to 4 of 4
Service Class Name Description Workload Name Base Resource Group Cpu Critical I/O Priority Group Honor Priority	<pre>. : FAST  <u>Velocity=80 goal</u>  <u>STCWORK</u> (name or ?)  <u>NO</u> (name or ?)  <u>NO</u> (YES or NO)  <u>NORMAL</u> (NORMAL or HIGH)  <u>NO</u> (DEFAULT or NO)</pre>	
Specify BASE GOAL informat E=Edit period, D=Delete per Period Action # Duration Imp.	ion. Action Codes: I=Insert new perio riod. Goal	
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Average response time of 00:00:01.0 Average response time of 00:00:10.0 Execution velocity of 80 ***** Bottom of data **********	00 00 *****

- Specifies whether work in this service class is exempted from default IFAHONORPRIORITY and IIPHONORPRIORITY processing
- Also for Service Class Overrides

#### **Recommendations for Honor Priority**

IBM

- Some zIIP work may be very latency sensitive and require to be dispatched quickly .
  - Namely some DB2 work, such as prefetch SRBs.
  - zIIP capacity may be constrained but CP capacity might be available to help
- Recommendation:
  - At the system level (IEAOPTxx) specify or default to IIPHonorPriority=Yes to allow CPs to help zIIP work.
  - Use the service class specific HonorPriority=No to selectively exclude work from receiving help.
    - Examples could be SPARK or Java batch that you do not want to be processed on general purpose processors

<ul> <li>Service Class (Period) header section of RMF Postprocessor Workload Activity report enhanced to provide 'HONOR PRIORITY' setting as defined in WLM service class definition</li> </ul>															
REPORT	BY: POLI	CY=STANDARD	WORKLOAD=STC	:	SERVIC CRITIC DESCRI	E CLAS AL PTION	SS=DB2P =NONE =DB2 subs	ystem	RESOURCE	GROUP=RO ORITY=NO	SPIX1		Prohit to re	oit offload egular CF	ling S
TRANCA	CTTONS	TRANC TIME			т /о	с г		CEDV			0/		MOTED	сто	DACE
AVG	1.00		0	SSCHRT	0.0	TOC	0		42.328	CP	4.70	BIK	0.000	AVG	1498.00
MPI	1.00	EXECUTION	0	RESP	0.0	CPU	33694ĸ	SRB	0.005	AAPCP	0.00	FNO	0.000	TOTAL	1492.94
ENDED	0	OUEUED	0	CONN	0.0	MSO	1233K	RCT	0.000	IIPCP	0.00	CRM	0.000	SHARED	1.00
END/S	0.00	R/S AFFIN	0	DISC	0.0	SRB	4204	IIT	0.000			LCK	0.002		
#SWAPS	3	INELIGIBLE	0	Q+PEND	0.0	тот	34931ĸ	HST	0.000	AAP	N/A	SUP	0.000	-PAGE-I	N RATES-
EXCTD	0	CONVERSION	0	IOSQ	0.0	/SEC	38812	AAP	N/A	IIP	0.00			SINGLE	0.0
AVG ENC	0.00	STD DEV	0			-		IIP	0.000					BLOCK	0.0
REM ENC	0.00					ABSRF	тν 39к							SHARED	0.0
MS ENC	0.00					TRX S	SERV 39K							HSP	0.0
TRANSAC	TRANSACTION APPL% : TOTAL : CP 4.70 AAP/IIP ON CP 0.00 AAP/IIP 0.00														

0.00

AAP/IIP

0.00

#### RMF enhancements for Spark

0.00

AAP/IIP ON CP

MOBILE : CP

IBM

#### Memory Limit for Resource Groups



- Specifies the maximum amount of memory that address spaces associated with the resource group through classification may consume on the local system (System Scope)
- The attribute is specified as absolute value in GB in the range 1 99,999,999.
- → IBM recommends that you use memory pools when it is required to limit memory consumption for new workloads such as Apache SPARK that provided guidance on how to operate them in a memory pool.



#### RMF enhancements for Spark ...

RMF PP Workload Activity report:	WORK	LOAD ACTI	VITY			
Service Policy Page	SYSDPLEX ION V2R2 RMF	START 02/10/2017-1 END 02/10/2017-1	.3.59.34 INTE .4.14.34	ERVAL 000.1	15.00 MODE = GOA	L
	POLICY ACTIVATION - SER	DATE/TIME 01/13/2 VICE POLICY PAGE -	2017 12.04.3	}	Memo loca	ry Limit on I system
INSTALL DATE: 01/13/2017 12.04.27 INS POLICY: STANDARD Standard policy I/O PRIORITY MANAGEMENT: YES DYNAMIC ALIAS MANAGEMENT: YES	TALLED BY: BMAI	-SERVICE IOC	DEFINITION CPU	SRB N	NISNC (O)SR MSO	000
SYSTEMS ID OPT SU/SEC CAP%TIME SYSD 00 79602.0 100 13.59.34 RESOURCE GROUPS	INTERVALID 00.15.00 SYSE	OPT SU/SEC 00 79602.0	CAP%TIME- 100 13.59.3	INTERVA 34 00.14.5	AL 59	
NAMEDESCRIPTION	SERVICE- CLASS	ACTUAL SUS CONSUMED	MIN	CAPACI MAX	ITY DEFINED AS	MEMORY- LIMIT
HIGHPRTY Very high priority work	STCSYS	3330 3330	5000	999999	SERVICE UNITS	1000
RGROUP1 resource group 1	GPMSERVE RMF RMFGAT SOAKER	830 200 440 190 0	0.75	11.25	NUMBER OF CPS	248

#### WLM shorter Response Time Goals

- The WLM Administrative Application is enhanced to allow the definition of service period response time goals below 15 milliseconds which is the minimum response time goal with z/OS V2R2 and below. The new minimum goal value is 0.001 seconds (one millisecond) and can be defined for base goals or when overriding attributes for a service class.
- When specifying an average response time goal, the total response time can be between 0.001 seconds (one millisecond) and 24 hours.

IWMAP34 A	verage response time goal	IWMAP35 Response time with percentile goal				
Enter a response ti	ne of up to 24 hours for period 2	Enter a percentile and response time goal for period 3				
Hours <u>00</u> Minutes <u>00</u> Seconds	(0-24) (0-99) MAS (0-9999)	Percentile <u>90</u> (1-99) Hours 00 (0-24)				
Importance <u>2</u> Duration 100	(1=highest, 5=lowest) 30 (1-999.999.999. or	Minutes <u>00</u> .(0-99) Seconds <mark>00.008</mark> (0-9999)				
	none for last period)	Importance <u>2</u> (1=highest, 5=lowest) Duration <u>10000</u> (1-999,999,999, or none for last period)				
F1=Help F2=Sp	lit F5=KeysHelp F9=Swap F12=Cancel					
		F1=Help F2=Split F5=KeysHelp F9=Swap F12=Cancel				

#### Shorter RT Goals – RMF Workload Activity Report



36

#### Shorter RT Goals – RMF Mon III SYSSUM

				DME	11000					DIMAD			6 0 1		
				RUL	V2R3	sys	plex Su	mmary	- UIC	PLXCB		_1ne 19	ot 34	Response time go	al.
WIM Same	les	: 4	79	Sus	tems: 8	3 Da	te: 02/	09/17	Time:	12.00	00 Rar	nde: 12	0 Sec	actuals and avera	<u>л</u> о
				-9-				1967) A 1							yc
				>>	>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	< XXXXX	xxxxxx	XXXXX	<mark>×</mark> <<<<	111				response times a	re
2532 (21										25 BITH 500				displayed in	
Service	Def	ini	tion:	SYST	EST			Insta	illed a	at: 11/1	15/16,	08.18.	14	milliseconds	
Act	ive	Ρc	licy:	BASE	POL			Activ	ated a	at: 02/0	09/17,	09.00.0	02	miniscoorida	
				2						<u>, 194</u>		-			
				G	oals ve	ersus	Actual	s		Irans	-Avg.	. Resp.	lime		
			Exec	Vel	Re	≥spon	se lime		Pert	Ended	WAII	EXECUT	ACTUAL		
Name		1	Goal	Act	Goa	al	Actu	al	Indx	Rate	lime	lime	lime		
OTOUT	0	5	0F	60		¥.			0 50	0.000					
STOLOU	5	4	30	69					0.50	0.000					
STULUW	3	3	25	40					0.56	0.192	040.0	400 E	1105		
SYSTER	W		117.0	81	1110					0.192	812.3	489.5	1165		
SYSSIC	5		NZH	64	NZH					0.192	0.000	0.000	0.000		
SYSIEM	3		NZA	84	NZA					0.000	0.000	0.000	0.000		
ISO_WLD	W			22						5.175	3.502	10395	10399		
TSOHIGH	S			67		and the second		100 Carl # 100		0.092	197.7	324.8	522.5		
		1		60	400	90%	1201210-0-0012-0-025	67%	4.00	0.075	0.000	284.5	284.5		
	2	2		100	3000	AVG	1593	AVG	0.53	0.017	1087	506.1	1593		
TSOLOW	S			19						5.083	0.000	10577	10577		
	1	3		27	500	85%		21%	****	3.583	0.000	8396	8396		
	2	3		7.1	5000	AVG	15786	AVG	3.16	1.500	0.000	15786	15786		
CONSOLES	R			40					NZA	0.000	0.000	0.000	0.000		
IXGLOGR	R			82					N/A	0.000	0.000	0.000	0.000		
TCPIP	R			20				2	NZA	0.000	0.000	0.000	0.000		
XCFAS	R			32			$\sim$		NZA	0.000	0.000	0.000	0.000		

TBM

#### WLM support for Container Pricing

#### IBM

Why new z/OS service definition objects?

Cloud workload paradigm asks for new ways of metering workloads in multi-tenant environments

IBM Z business asks for an infrastructure to support novel pricing options  A Tenant Report Class (TRC) is similar to a WLM Report Class.
 TRCs are assigned through WLM classification and are always associated with a Tenant Resource Group.  A Tenant Resource Group (TRG) is somewhat similar to a WLM Resource Group and can be associated with tenants or solutions.

 TRGs aggregate consumption data and can optionally be used to apply consumption limits.

## WLM/SRM enhancements for Container Pricing

- New panels for defining and modifying Tenant Resource Groups
- New panels for defining and modifying Tenant Report Classes
- Resource Group (RG) and Tenant Resource Group (TRG) enhancements
  - New Sysplex-wide Type 4 limit expressed at a scale of "MSU"
  - · Optionally, specialty processor consumption can be counted towards the limit
- A new service definition option that allows to disable "Discretionary Goal Management" globally
- WLM/SRM programming services are enhanced
  - A new IWM4QTNT service allows monitoring products to retrieve TRG-level consumption data
  - Various APIs provide TRC and TRG indications
- WLM TRG level data will be reported by monitoring products, such as RMF
  - Also in SMF type 70 record
- SMF70 TRG data can be consumed in SCRT
- Various other components to be updates (SMF, SDSF, z/OSMF, ...)

#### **Container Pricing workflow overview**

- For a eligible solution, IBM provide customers with a "Solution ID" (key)
- In their WLM service definition customers
  - Define one or more TRGs and paste the Solution ID into the definition
    - Dummy solution IDs for test/education will be documented
  - Define one or more TRCs associated with TRG
  - Change or add classification rules to classify eligible work (only!) and assign service class and TRCs
  - Install and activate WLM service definition
- Monitors query WLM and write new SMF70 data sections for TRGs
  - In addition, the TRC and TRG data will be reported via the existing report class and resource group mechanism in the WLMGL report
- SCRT consumes SMF70 and SMF89 data for billing
  - Verifies solution ID, applies pricing rules





#### TRG and TRC definition

- The WLM Administrative Application Level increased to 32.
- Tenant Resource Groups and Tenant Report Classes can be defined via new menu items.
- Specification of these new objects will increase the functionality level of the service definition to 32.

Functionality LEVEL032 Definition Menu WLM Appl LEVEL032 Definition data set . . : 'WLM.DEMO.SRVDEF.XML' Definition name . . . . . PROD01 (Required) Description . . . . . . Production service definition Select one of the following options. 1. Policies 12. Tenant Resource Groups Workloads 13. Tenant Report Classes 3. Resource Groups 4. Service Classes 5. Classification Groups 6. Classification Rules Report Classes Service Coefficients/Options Application Environments 10. Scheduling Environments 11. Guest Platform Mgmt Provider

#### IBM

#### Tenant Resource Group (TRG) definition

- The TRG name is mandatory (8 char)
- Description, Tenant ID, Tenant Name are optional and are expected to be used in a z/OS cloud context
- For qualified offerings, a 64 char Solution ID needs to be provided.
- Exactly enter (paste) the IBM provided Solution ID string
- WLM performs sanity check only.
- Solution ID considered during SCRT processing. Multiple TRGs may specify same Solution ID
- <u>TRG capacity limits</u> should not be specified unless there is a need to limit processor consumption. There is NO minimum consumption limit. There is NO memory limit.

Create a Tenant Resource Group
Enter or change the following information: Tenant Resource Group Name TRGDEM01 (required)
Description Sample TRG
Solution ID: ZS1ZZZA-I1F9FAD-6B20999D15-85370C8D-726F-458C-B84A-37D01D-A10E44
Tenant ID
Tenant Name Solution Newapp
Define Capacity: 1. In Service Units (Sysplex Scope) 2. As Percentage of the LPAR share (System Scope) 3. As a Number of CPs times 100 (System Scope) 4. In accounted workload MSU (Sysplex Scope)
Maximum Capacity
Include Specialty Processor Consumption NO (YES or NO)

#### Tenant Report Class (TRC) definition

- The TRC name is mandatory (8 char)
  - Name must be unique (also across report classes)
  - Up to 2047 Report Classes and Tenant Report Classes can be defined
- The TRG name is required, i.e. any TRC must be associated with a TRG
- Monitoring interfaces and monitors report on TRC as on standard report classes
- Highly recommended: When assigning TRC in classification rules, do only use homogeneous TRCs, i.e. do not mix of service classes for a given TRC.

Create a Tenant Report Class
inter or change the following information:
enant Report Class Name T_CDC (Required) Description
enant Resource Group Name TRGDEM01 (Required; name or ?)

## Comparison of TRG and Resource Group Capping Types

Туре 1	Туре 2	Туре 3	Туре 4					
Raw CPU+SRB service units ("Raw" meaning that Service Definition Coefficients are not applied)	Percent of CP LPAR share (even if specialty processor consumption included). May exceed 100%.	Percent of one CP processor (even if specialty processor consumption included)	Processor consumption expressed in "accounted workload MSU"					
Limit applies to Sysplex	Limit applies to each System	Limit applies to each system	Limit applies to Sysplex					
For <b>all</b> (T)RG types only captured TCB and SRB times are counted towards the limit. The limit is enforced based on a one minute average (i.e., no 4HRA). Up to 32 RGs <b>plus 32 TRGs</b> may be defined.								

IBM

# SMF Updates



## What Are High Frequency Throughput Statistics?

- IBM providing very granular SMF records for diagnosing performance issues
  - 5 to 60 second interval
- New SMF record defined which will hold this performance data
- Can help diagnose short running but repeat problem that otherwise may not be measurable with current monitors and reports
  - Top/bottom of the hour
  - Spikiness in arrival rate
  - Customer reports periodic/sporadic issues, RMF shows no issues
- New records have short interval and precise insights into system performance

#### Value of SMF 98 Records

- Subtype 1 is first supported subtype
  - Information on supervisor component about the workload and it's significant jobs
- Currently available information:
  - CPU Utilization
  - System spin locks and most contentious spin locks
  - Suspend Locks
  - Summary of the workload's CPU delay by CPU type and dispatch priority group
    - high, medium, low, and discretionary, via. WLM priorities
  - Consumption sections
    - CPU utilization data organized by work unit type, execution efficiency data by thread density, and spin lock contention data. A consumption section is divided into subsections by CPU type (except for spin lock data), by dispatch priority group, and by relative CPU utilization. Each consumption subsection contains both summary data from all contributing address spaces and data from the most significant address space



## **Enabling High Frequency Throughput Statistics**



- Enabled via APAR OA48570 and OA48571
- Provides data points only seconds apart
- High Frequency Throughput Statistics (HFTS) Written to SMF 98 record
- Writing of SMF records controlled completely in SMFPRMxx member of PARMLIB

#### Collecting HFTS SMF 98 Records - SMFPRMxx

- Type statement must include the new 98 SMF record type
  - TYPE(0,23,30,41,42,70:79,80:83,88:90,**98**,99(1,2,6,14), 100:103,113,120))
- HFTSINTVL must be specified with timing of records
  - Supported values of 5, 10, 15, 20, 30, and 60 seconds
  - Example: HFTSINTVL(20) for 20 second interval
- NOHFTSINTVL Disable HFTS (default value if no HFTSINTVL specified)
  - Specifying NOHFTSINTVL but leaving TYPE(98) will prevent writing of records

#### HFTS recommendations

- Set initial HFTSINTVL value to 20
  - HFTSINTVL(20)
- · Good balance of small interval and with amount of overhead
  - Consider lowering to 5 for problem period diagnosis
  - Raise to 60 to minimize overhead and record volume if steady state has few or no anomalies to continuously diagnose
- Make use of logstream recording of SMF data (Are you using logstream for SMF?)
  - Put SMF 98 records to their own logstream
  - For Example: LSNAME(IFASMF.SYS.SMF98.DATA),TYPE(98))
- Isolation of 98 records will limit any performance impact to other record types
  - Will also keep other logs consistent if starting/stopping HFTS process

#### z/OS SMF Real-time Services

- New Real-Time SMF services provided on top of existing buffer technology
- Define new "In-memory (INMEM) Resources" for specific records
- Can write SMF records to a real-time resource only No disk required
  - Co-exists with current SMF logstream technology
  - Not supported when RECORDING(DATASET) used
- APIs allow application to access SMF data as it is buffered
  - Unauthorized access policed via SAF
  - Connect/Get/Disconnect model similar to traditional QSAM access
- Potential use-cases include
  - Detecting security violations in real-time
  - Real time monitoring resource usage
  - Dynamic Job Scheduling based on current resource consumption



#### SMF callable services overview

- New SMF callable services can provide bulk in-memory data retrieval or real-time data retrieval.
- Historical data, when available, can be pulled from in-memory buffer. When not available the data can be obtained from traditional longterm storage however the cross-over is NOT transparent, the application needs to manage this.



#### **Define an In-memory Resource**

- Up to 32 in-memory resources can be defined
  - Each resource can have a unique set of records collected to be read
  - Up to 8 programs can connect to resource to read records
  - Resource name similar to LSNAME IFASMF.xxxx up to 26 characters
- RESSIZMAX can be used to define the buffer size
  - Note: In memory buffer treated as a wrapping buffer that is never emptied
- Accepts all expected TYPE or NOTYPE statements
  - The TYPEs recorded in-memory will not be processed by DEFAULTLSNAME processing
    - Allows for InMemory only records (i.e. never written to permanent storage)



#### Putting the pieces together

- How can this new data access pattern be used?
  - Turn on for any SMF records required for real-time analysis
  - Turn on SMF records that are "transient" They do not require long-term storage and are only valuable for real-time analysis
  - Control access to real-time data Only give SAF access to specific users
  - No need to change existing SMF configuration Continue to store long-term data in data sets or logstreams





#### **Use Cases SMF Real-Time Analytics**

Detect excessive memory consumption – SMF 30

 Monitor high water mark for real memory usage for jobs and send alerts if usage exceeds normal consumption
 INMEM (IFASMF.INMEM.RES5, TYPE (30), RESSIZMAX (128M))

#### • Detect security violations in real-time – SMF 80

 Monitor volume of datasets/files accessed per user within a given time period and raise alerts for above normal access rates

INMEM(IFASMF.INMEM.RES5,TYPE(80),RESSIZMAX(128M))



#### **RMF Information and Tools**

- website: www.ibm.com/systems/z/os/zos/features/rmf/
  - Product information, newsletters, presentations, ...
  - Downloads
    - Spreadsheet Reporter
    - Postprocessor XML Toolkit
- RMF email address: rmf@de.ibm.com



- Documentation and news:
  - RMF Report Analysis, SC34-2665
  - RMF User's Guide, SC34-2664
  - RMF Programmer's Guide, SC34-2667
  - Latest version of PDF files can be downloaded from: www.ibm.com/servers/resourcelink/svc00100.nsf/pages/zOSV2R3RmfPublications?OpenDocument



56

#### **RMF** Function Reference

Function Availability RMF z14 Support RMF support for Virtual Flash Memory **APAR OA50761** RMF support for 10GbE RoCE Express2 card **APAR OA50762** RMF support for Crypto Express6S (CEX6) card **APAR OA50693** RMF support for zHyperLink APAR OA50755 RMF z14 Toleration support APAR OA51913 RMF support for SuperPAV APAR OA49415 RMF support for IBM zHyperWrite z/OS 2.3 (Coexistence support OA48870) RMF Asynchronous CF duplexing for lock structures support **APAR OA49148** RMF Enhancements for CF Encryption APAR OA52003 RMF Enhancements for Spark **APAR OA50760** RMF Support for WLM Shorter Response Time Goals z/OS 2.3 RMF support for 2GB Large Pages APAR OA48913

## z/OS Workload Management – More Information



IRM

z/OS WLM Homepage:

http://www.ibm.com/systems/z/os/zos/features/wlm/ - Inside WLM: https://ibm.biz/BdF4L4

- z/OS MVS documentation
  - z/OS MVS Planning: Workload Management: <u>http://publibz.boulder.ibm.com/epubs/pdf/iea3w101.pdf</u>
  - z/OS MVS Programming: Workload Management Services: <u>http://publibz.boulder.ibm.com/epubs/pdf/iea3w201.pdf</u>
- IBM Redbooks publications:
  - System Programmer's Guide to: Workload Manager: <u>http://publib-b.boulder.ibm.com/abstracts/sg246472.html?Open</u>
  - ABCs of z/OS System Programming Volume 12
     <u>http://publib-b.boulder.ibm.com/abstracts/sg247621.html?Open</u>

