

Thomas Blattmann

z/OS Workload Management Development, Boeblingen, Germany

blattmann@de.ibm.com



Workload Management Update for z/OS 1.10 and 1.11



Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

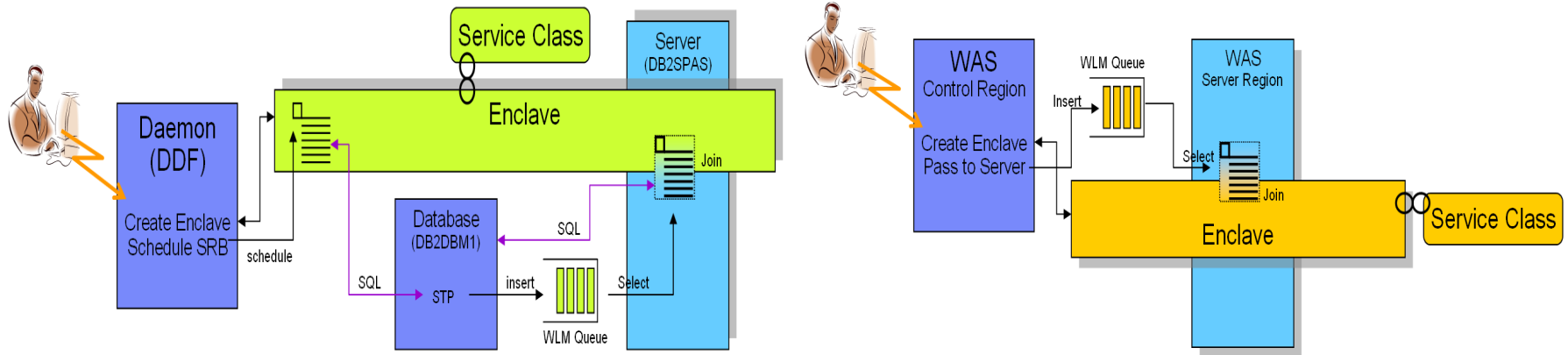
Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Agenda

- Enclave Enhancements
 - Enclave Server Management
 - Work-Dependent Enclaves
- WLM Management
 - LDAP Support
 - Resource Group Enhancements
 - Do not always honor Skip Clock in Policy Adjustment
- WLM Reporting
 - Extend Number of Report Classes
 - Additional Group Capacity Information in RMF
- Externalized OPT Information
- Enhanced Storage Monitoring
- WLM Tools Overview

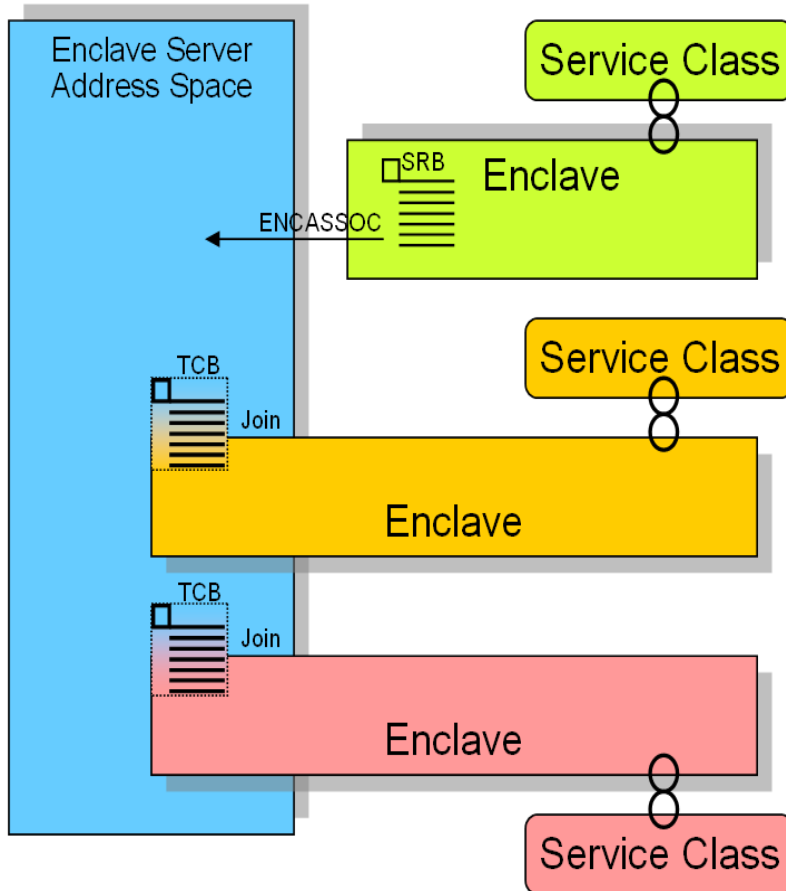
WLM Enclaves – An Overview



- An **enclave** is a transaction that can span multiple dispatchable units (SRBs and tasks) in one or several address spaces and is reported on and managed as one unit.
- The enclave is managed separately from the address spaces it runs in.
 - CPU and I/O resources associated with processing the transaction represented by the enclave are managed by the transaction's performance goal.
 - Storage (MPL level, paging) of the address space is managed to meet the goals of the enclaves it serves (if enclave server address space) or to the performance goal of the address space (if no server address space).

WLM Enclave Server Address Spaces

A Short Retrospective



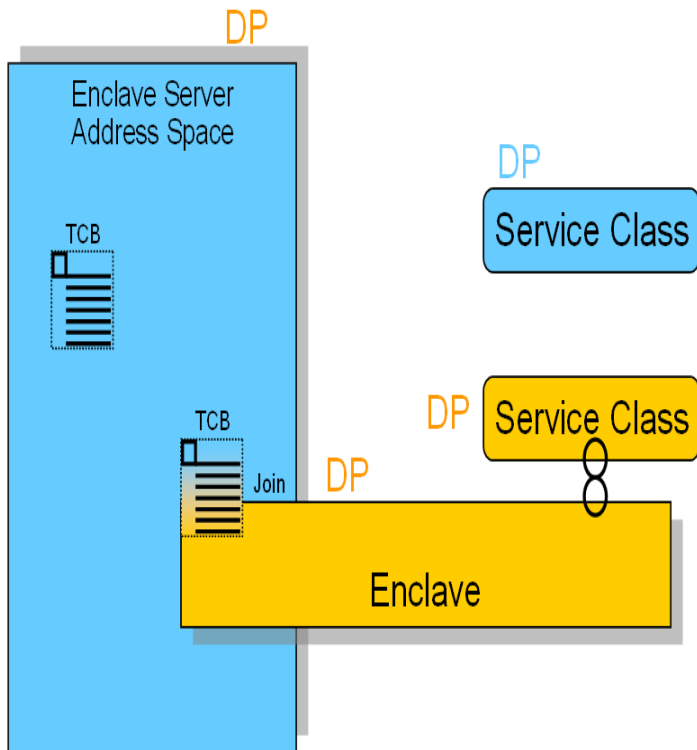
- An address space becomes an enclave server when
 - An enclave SRB issues SYSEVENT ENCASSOC
 - A TCB of the address space joins an enclave, and does not specify ENCLAVESERVER=NO (which is typically not the case)

- Assumption (Programming Model)
 - All work being executed within the address space is related to enclaves
 - That means
 - There is no significant amount of work (TCBs) executing in such address spaces which is not related to enclaves

- Enclave Server Management
 - CPU and I/O DP is derived from service class of most important enclave
 - Meaning: No CPU and I/O management exists for these server address spaces
 - Storage management is done to meet the served enclave's goals.

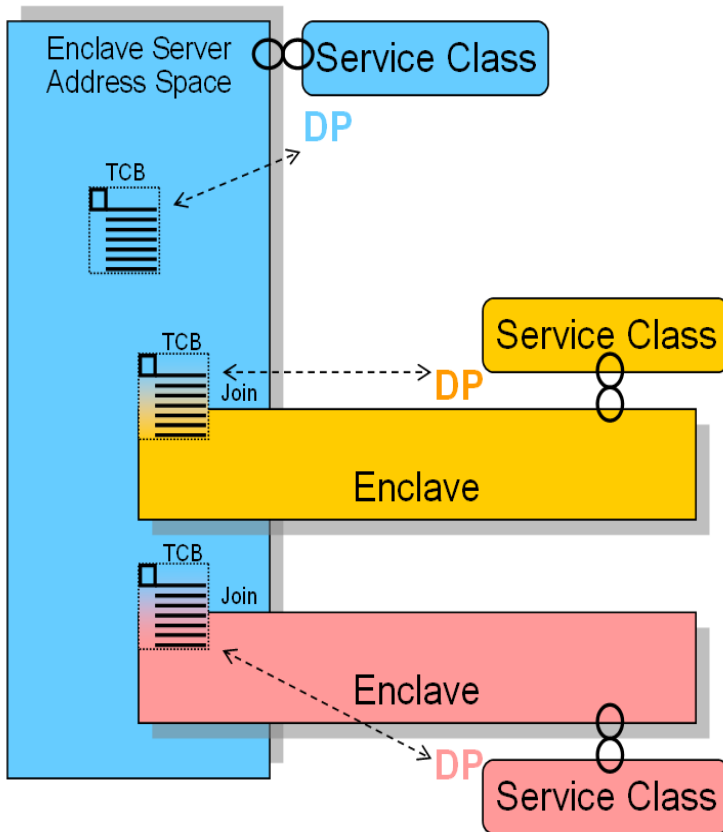
WLM Enclave Server Management

Is There a Possible Problem?



- What if the programming model does not hold true?
 - What happens if there is significant work running in TCBs not associated with enclaves?
 - Example: Garbage collection for a JVM (WAS)
 - Example: Common routines which provide service for the enclave TCBs
 - Is it sufficient to manage this work in the same way as the enclaves?
- What happens if no enclaves are running in server address spaces ?? (this applies to queue servers only)
 - And the address space is swapped out?
 - A mechanism exists to swap in the address space but this mechanism assumes that the swap in is only for a queue server task which wants to select a unit of work and then joins the enclave. If no enclave is joined, the address space is again swapped out.
 - And even if the address space stays swapped in?
 - The TCBs running within the address space just stay with the DP and IOP from the last enclave being associated with the address space.
 - No CPU or I/O adjustment is performed.

WLM Enclave Server Management Changes with z/OS 1.12



- **New OPT Parameter**
 - ManageNonEnclaveWork = {No|Yes}
 - Default: No (no change to previous releases)
 - Causes everything in the address space, which is not associated to an enclave, to be managed towards the goals of the external Service Class to which the address space has been classified to.

- **Advantages**
 - Enclave (Queue) server address spaces in which no enclave is running will be managed as usual address spaces.
 - The importance and goal of the service class for the address space now has a meaning.

- **Attention**
 - **The importance and goal of the service class for the address space now has a meaning**
 - Therefore verify goal settings for server address spaces
 - This is a deviation from the past when the service class for servers was only important for startup, shutdown and recovers

Work-Dependent Enclaves

- Background

zIIPs allow middleware components to run a certain percentage of their work “offloaded” from regular processors.

The offload percentage is an attribute of the enclave under which the unit of work runs.

The offload percentage is defined by the middleware component via a (not generally published) WLM interface.

- Limitations

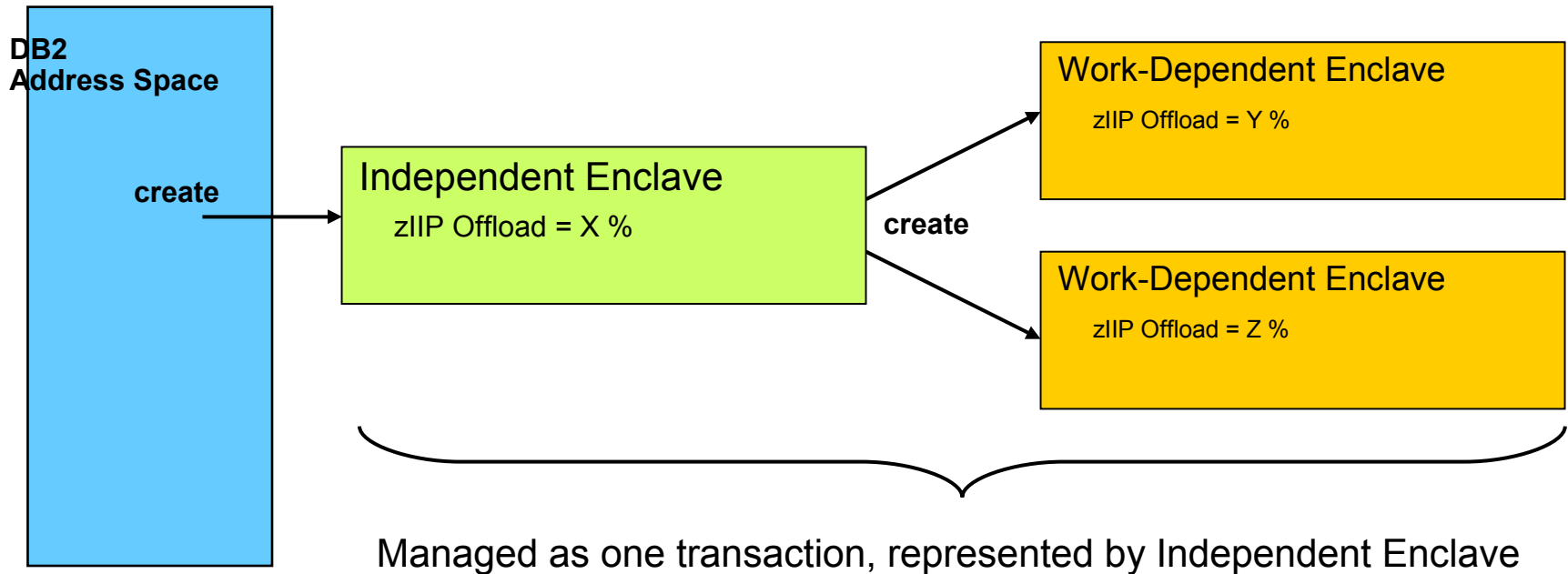
It is not possible to specify different offload percentages for different units of work running under the same enclave.

- Intended Use Case

DB2/DDF wants to specify different offload percentages for the different units of work of a parallel query,

AND still wants to maintain the transactional context to run the units of work under the same “SRM Transaction” (enclave).

Work-Dependent Enclaves



▪ Solution

Implement a new type of enclave named “Work-Dependent” as an extension of an Independent Enclave. A Work-Dependent enclave becomes part of the Independent Enclave’s transaction but allows to have its own set of attributes (including zIIP offload percentage).

Work-Dependent Enclaves

SDSF Enclave Panel

```

  Display  Filter  View  Print  Options  Search  Help
-----
SDSF ENCLAVE DISPLAY  SYS1  ALL  LINE 1-8 (8)
COMMAND INPUT ==>
PREFIX=*  DEST=(ALL)  OWNER=*  SYSNAME=SYS1  SCROLL ==> CSR
NP  NAME  Status  Type  SrvClass  Per  RptClass  CPU-Time  OwnerAS  Re
28000000006  ACTIVE  IND  VEL_1  1  RC_2  0.00  36
2C000000008  ACTIVE  WDEP  VEL_1  1  RC_2  0.83  36
30000000007  ACTIVE  WDEP  VEL_1  1  RC_2  0.83  36
34000000009  ACTIVE  WDEP  VEL_1  1  RC_2  0.83  36
3800000000A  ACTIVE  WDEP  VEL_1  1  RC_2  0.83  36
3C00000000B  ACTIVE  WDEP  VEL_1  1  RC_2  0.83  36
24000000002  INACTIVE  DEP  SYSSTC  1  RC_0  0.00  22
20000000001  INACTIVE  DEP  SYSTEM  1  RC_0  0.00  7

F1=help  F2=SPLIT  F3=END  F4=f 'SYSM0  F5=IFIND  F6=FIND '-'
F7=UP  F8=DOWN  F9=SWAP nex  F10=LEFT  F11=RIGHT  F12=cRETRIEV
MA c 05/021
  
```

Work-Dependent Enclaves

RMF Monitor III

```

RMF V1R12 Enclave Report
Command ==>
Line 1 of 7
Scroll ==> CSR
Samples: 100 System: SYS1 Date: 02/23/10 Time: 03.06.40 Range: 100 Sec
Current options: Subsystem Type: ALL -- CPU Util --
                  Enclave Owner: Appl% EAppl%
                  Class/Group: 0.1 1.4
Enclave Attribute CLS/GRP P Goal % D X EAppl% TCPU USG DLY IDL
*SUMMARY
ENC00006 VEL_1 1 5 W 0.812
ENC00002 VEL_1 1 5 W 0.163 2.530 1.0 0.0 0.0
ENC00004 VEL_1 1 5 W 0.162 2.528 0.0 0.0 0.0
ENC00005 VEL_1 1 5 W 0.162 2.519 0.0 0.0 0.0
ENC00003 VEL_1 1 5 W 0.162 2.518 0.0 0.0 0.0
ENC00001 VEL_1 1 5 W 0.000 0.007 0.0 0.0 0.0
F1=HELP F2=SPLIT F3=END F4=RETURN F5=RFIND F6=TOGGLE
F7=UP F8=DOWN F9=SWAP lis F10=BREF F11=FREF F12=RETRIEVE
MA c 03/015

```

Enclave Enhancements: Availability

Function	z/OS V1.12	z/OS V1.11	z/OS V1.10	Older Releases
Non Shell Server Management	+			
Work-dependent Enclaves	+	+	OA26104	OA26104 → z/OS 1.8

- Non Shell Server Management

New OPT Parameter ManageNonEnclaveWork=YES/NO. Default is NO, meaning the function is not yet enabled.

- Work-Dependent Enclaves

- New function available with WLM APAR OA26104
- DB2 exploitation with APAR PK76676
- SDSF support with APAR PK74125
- RMF support with z/OS 1.11

WLM Management: LDAP Subsystem is supported

	LDAP
Accounting Information	
Collection Name	
Connection Type	
Correlation Information	
EWLM Service Clas	
EWLM Transaction Class	
LU Name	
Netid	
Package Name	
Perform	
Plan Name	
Priority	
Procedure Name	
Process Name	
Scheduling Environment Name	
Subsystem Collection Name	
Subsystem Instance	●
Subsystem Parameter	
Sysplex Name	●
System Name	
Transaction Class/Job Class	
Transaction Name/Job Name	●
Userid	

- Work requests include all work processed by the z/OS LDAP server.

- Supported Work Qualifiers

- Subsystem Instance (SI)

The z/OS LDAP server's job name. Needed to distinguish between different LDAP servers.

- Transaction Name/Job Name (TN)

The z/OS LDAP server's enclave transaction name. "GENERAL" for all LDAP work that is not assigned a user-defined exception class. Any transaction name that is also defined in the configuration file of the directory server.

- For further information see

z/OS IBM Tivoli Directory Server Administration and Use for z/OS (SC23-5191-XX)

WLM Management:

Subsystems supported by the WLM Administrative Application

	A S C H	C B	C I C S	D B 2	D D F	E W L M	I M S	I W B	J E S	L D A P	L S M	M Q	N E T V	O M V S	S M	S T C	T C P	T S O	S Y S H
Accounting Information	●			●	●				●					●		●		●	
Collection Name		●		●	●										●				
Connection Type				●	●														
Correlation Information				●	●														
EWLM Service Class						●													
EWLM Transaction Class						●													
LU Name			●	●	●		●						●						
Netid				●	●		●												
Package Name				●	●														
Perform				●	●				●							●		●	
Plan Name				●	●														
Priority				●	●				●			●	●						
Procedure Name				●	●							●							
Process Name				●	●							●							
Scheduling Environment Name				●	●				●										
Subsystem Collection Name				●	●				●										
Subsystem Instance		●	●	●	●		●	●	●	●	●	●	●				●		
Subsystem Parameter				●	●			●				●			●	●			
Sysplex Name	●	●	●	●	●		●	●	●	●	●	●	●	●	●	●		●	●
System Name	●													●		●		●	●
Transaction Class/Job Class	●	●		●			●	●	●		●	●							
Transaction Name/Job Name	●	●	●	●			●	●	●	●	●	●	●	●		●	●		
Userid	●	●	●	●	●		●	●	●	●	●	●	●	●	●	●		●	



Not relevant anymore

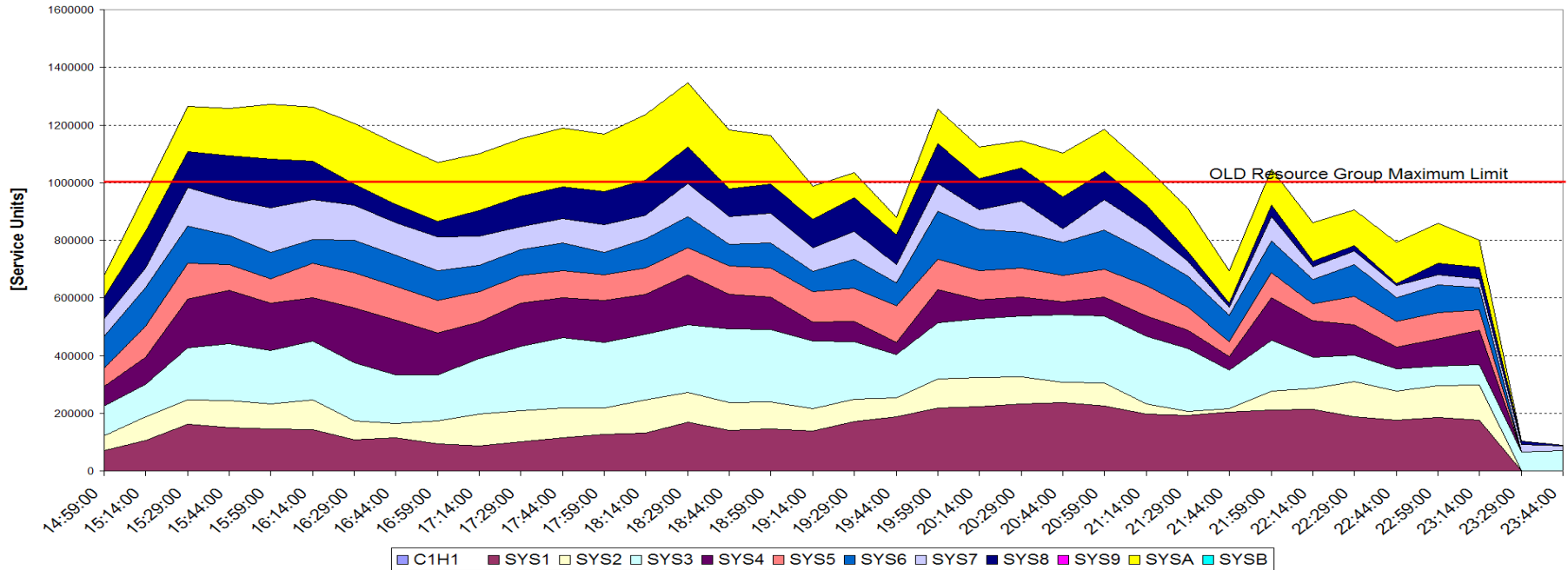


Latest supported subsystems

WLM Management: Resource Group Type 1 Limitations

Update Title

Service Unit Consumption for each System
Service Class: BATCHSTD Period: 1



- Type 1 Resource Groups provide sysplex-wide limits for CPU consumption.
- Prior to OA29704 (for z/OS 1.10 and 1.11) a minimum and maximum of 999,999 SUs/sec was the highest possible definition.

WLM Management: Resource Group Enhancements

```
vamp - wc3270
Resource-Group Xref Notes Options Help
-----
Command ==>
Modify a Resource Group

Enter or change the following information:

Resource Group Name . . . . . RGROUP1
Description . . . . . Res_Group_No_1

Define Capacity:
1_ 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)
Minimum Capacity . . . . . 0000
Maximum Capacity . . . . . 15000000
```

```
vamp - wc3270
File Utilities Notes Options Help
-
FC
Resource Group Compatibility Check

One or several Resource Group capacity values in this WLM Service
Definition exceed 999,999. Please make sure to have APAR OA29704
applied on all systems in your Sysplex that are currently at level
z/OS V1R11 (FMID HBB7760) or previous releases. If this APAR is not
installed, the Service Definition can get corrupted when it is
opened, modified, or installed via such a system and will not be
extractable from there.

Have you applied the necessary maintenance (OA29704) to all systems
running with z/OS V1R11 and previous releases in your Sysplex ?
1_ 1. YES, WLM APAR OA29704 has been installed on all systems
   2. NO, I am not sure if the necessary service is installed and
      I will have to check first. Terminate the installation of
      the Service Definition.

F1=Help      F2=Split     F5=KeysHelp  F9=Swap      F12=Cancel
```

- OA29704 for z/OS 1.10 and z/OS 1.11 allows you to specify new limits of up to 8 digits.
- Because this is a PTF (APAR) a warning message is shown when a min/max capacity value greater than 6 digits is entered.

Make sure that all systems are at least at z/OS 1.10 with OA29704 applied before installing and activating such a service definition

- On systems w/o this support the WLM Administrative Application is not able to extract the service definition from the Couple Data Set.
- On systems w/o this support the WLM Administrative Application would truncate the resource group capacity values to 6 digits if it is attempted to read the Service Definition from ISPF tables.

Regardless of whether or not the APAR has been applied, systems w/o the support honor the definition during runtime.

WLM Management: Do Not Always Honor Skip Clock



- What is the skip clock ?
 - If WLM can't help a service class it sets a skip clock to not assess it in the next 3 policy adjustment cycles.
 - This is done for efficiency reasons and to help other work.
- Is this always a good thing to do ?
 - Usually yes !!
 - As long as many service classes (10 or more) have been defined it is usually the case that more than 1 service class miss its goals
 - But
 - In the rare case that only a few service classes are defined in a service definition then also only 1 or 2 can miss their goals.
 - In this event it is not beneficial to no longer assess a service class for 3 consecutive policy adjustment cycles
 - Especially when it might be possible to help the work with IRD Weight Changes
 - In this event the situation on another LPAR can change and might make it possible to help a service class in the next policy adjustment cycle
- Solution introduced with z/OS 1.11

The skip clock will no longer be honored if 5 or less service class periods do not meet their performance objectives.




WLM Management Availability

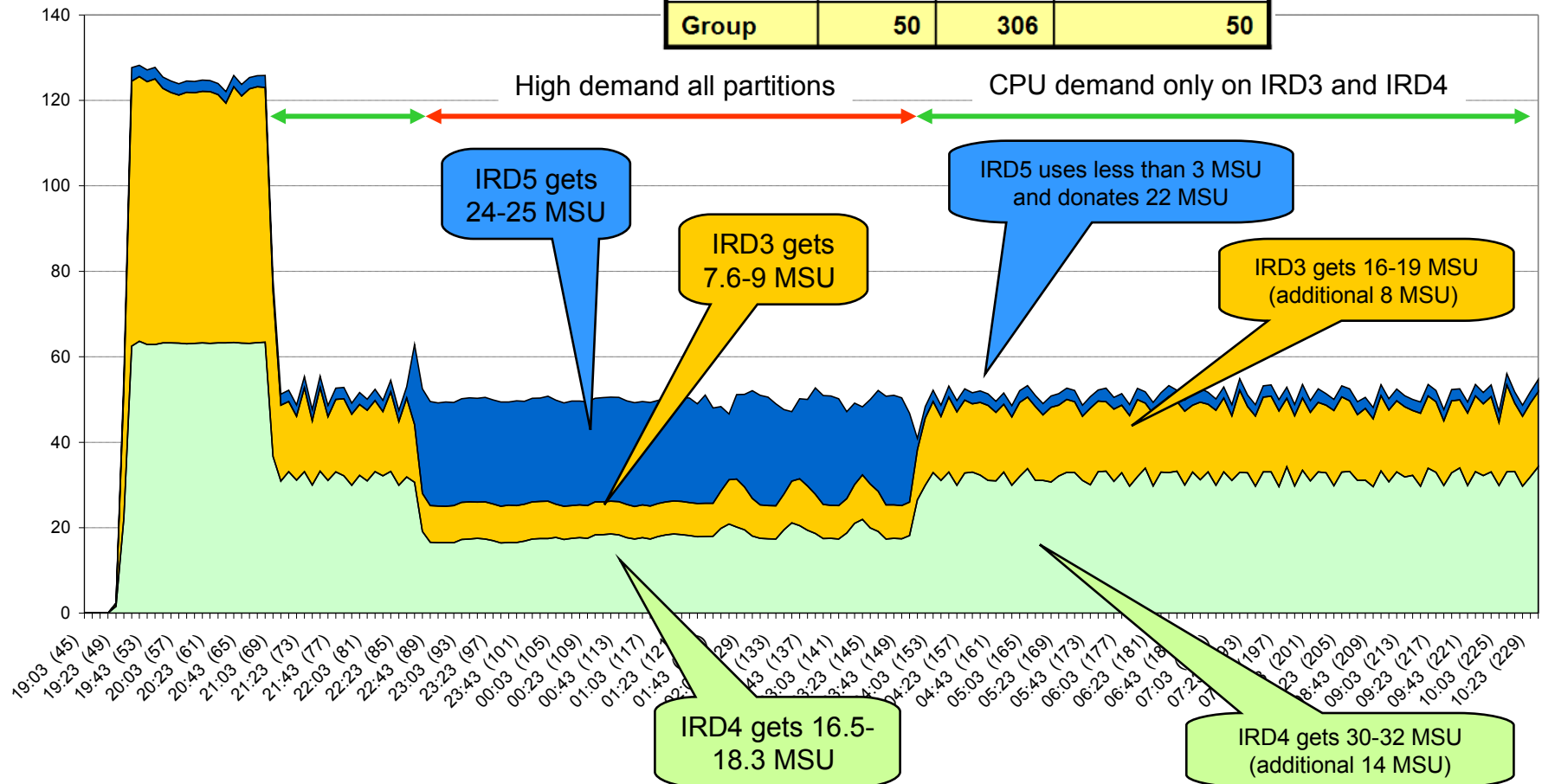
Function	z/OS V1.12	z/OS V1.11	z/OS V1.10	Older Releases
New Resource Groups (Type 2 and Type 3)	+	+	+	z/OS 1.8
8 digit resource group minimum and maximum (for Type 1)	+	OA29704	OA29704	
Change in skip clock processing	+	+		
LDAP Support	+	+		

Group Capacity: Summary

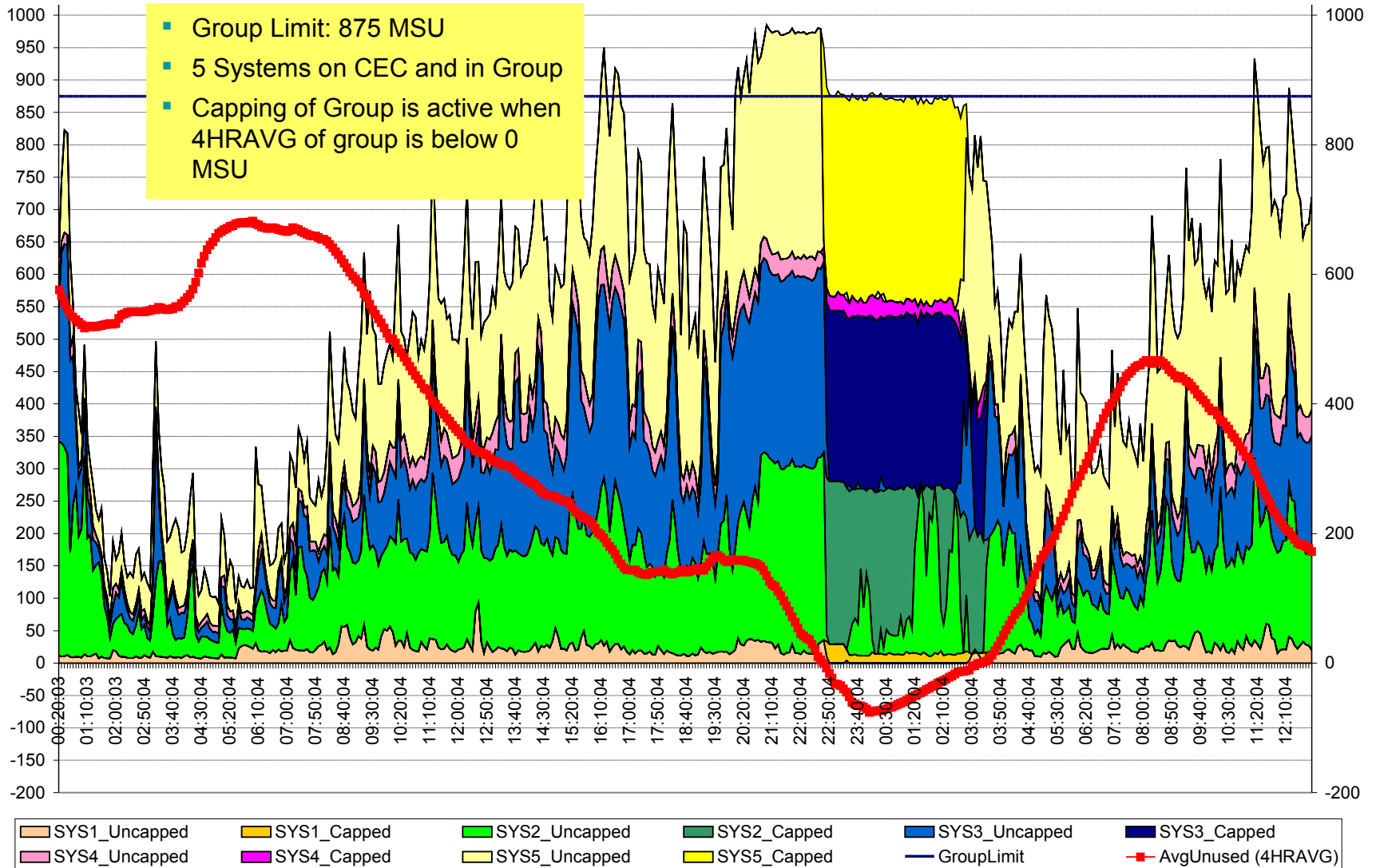
- Is based on defined capacity
 - Each partition obtains information for the other partitions of the group from PR/SM
 - Calculates the group consumption and whether the group should be capped
 - If the group becomes subject to capping
 - The partition calculates whether it is above or below of its entitlement
 - If it is above its entitlement the partition must apply capping (phantom weight or cap pattern)
- The entitlement of a partition is its share based on its weight within the group (named target MSU)
 - In addition if not all partitions use their entitlement the partition can obtain unused MSUs
 - The partition can always use its target MSU value assuming the overall LPAR definitions allow it
- Group Capacity and Defined Capacity can be combined
 - The z/OS system will always honor the smaller of both capacity limits
- It is possible to define multiple capacity groups on a CEC
 - A partition can only belong to one group
- Working with IRD CPU Weight Management
 - Defined and Group Capacity work with IRD but
 - Weight Changes are only possible for partitions which are not being capped (or subject to capping)
- Restrictions: Defined and Group Capacity
 - A partition must not be defined with dedicated processors
 - The partition must be defined with shared processors and WAIT Completion = NO
 - Initial Capping must not be defined
 - z/OS must not run as a VM guest
- PR/SM capping works within $\pm 3.6\%$ from the defined capping value

Group Capacity: Demo Scenario

Partition	Limit	Weight	Target MSU
IRD3 	n/a	52	~8.5
IRD4 	n/a	102	~16.7
IRD5 	n/a	152	~24.8
Group	50	306	50



Group Capacity: Customer Example

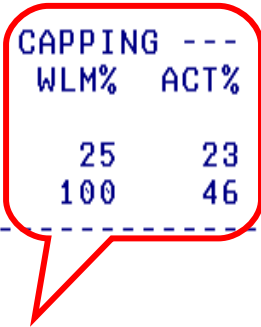


RMF z/OS 1.11 Enhancements for Group Capacity...

GROUP CAPACITY REPORT

z/OS V1R11 SYSTEM ID TRX1 DATE 02/26/2009 INTERVAL 05.00.000
 RPT VERSION V1R11 RMF TIME 11.00.00 CYCLE 1.000 SECONDS

GROUP-CAPACITY NAME	LIMIT	PARTITION	SYSTEM	-- MSU --		WGT	---	CAPPING		- ENTITLEMENT -	
				DEF	ACT			DEF	WLM%	ACT%	MINIMUM
RMFGRP	60	TRX1	TRX1	100	4	400	NO	25	23	40	60
	~	TRX2	TRX2	100	13	200	NO	100	46	20	60
-----				TOTAL		17	600				

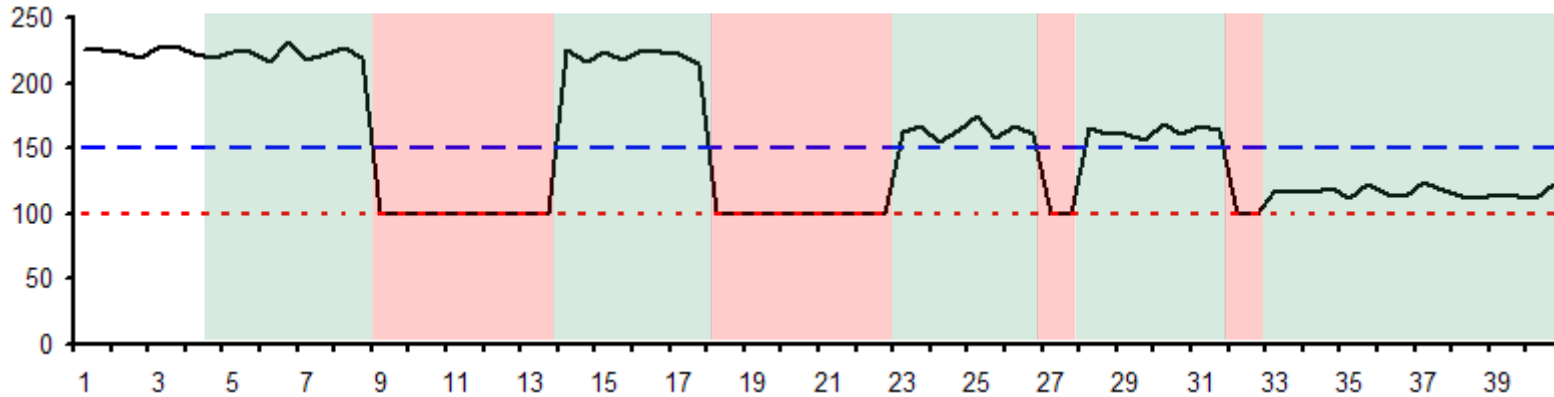


Field Heading	Meaning
CAPPING WLM%	Percentage of time when WLM considers to cap the partition
CAPPING ACT%	Percentage of time when capping actually limited the usage of processor resources for the partition

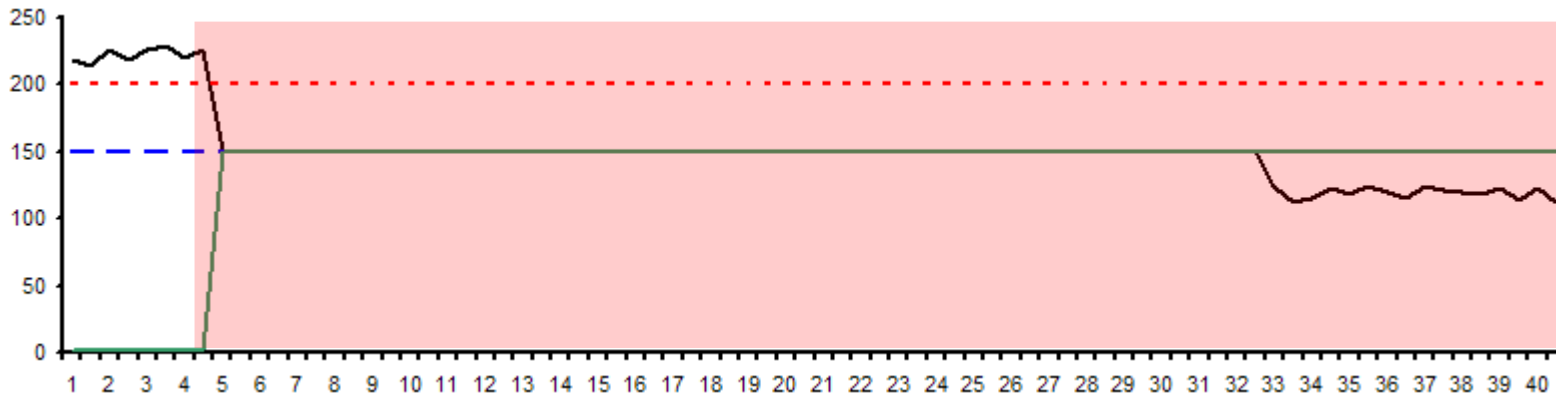
WLM Capping

Cap Pattern vs. Phantom Weight

Capping with **Cap Pattern** (when Soft Cap > MSU@LparWeight)



Capping using a **Phantom Weight** (when Soft Cap < MSU@LparWeight)



--- Soft Cap

..... MSU @ LPAR Weight

— MSU with Phantom Weight

RMF z/OS 1.11 Enhancements for Group Capacity...

Capping WLM% versus ACT%

- Capping WLM% = $SMF70NSW * 100 / SMF70DSA$
 - SMF70NSW is incremented for each Diagnose sample with the WLM-capped flag ON. The flag is ON if the LPAR was capped via Diagnose 0304.
- Capping ACT% = $SMF70NCA * 100 / SMF70DSA$
 - SMF70NCA is incremented for each Diagnose sample which indicates an actual-MSU-consumption below the MSU-at-weight factor
 - The pricing management adjustment weight of the LPAR (aka phantom weight) is added to the total of all active-logical-partition weights to compute the fraction of processor resources that the LPAR may use (MSUatWgt)
 - The actual MSU consumption of the LPAR is computed from the total dispatch time measured between two Diagnose samples
 - Following calculations done in RMF after Diagnose 0204 was called:

$$MSUatWeight = \frac{\text{Current LPAR weight} * \text{CPC capacity in MSUs}}{\text{Total weight} + \text{PMA weight}}$$

$$ActualMSU = \frac{\text{Dispatch time delta} * 3600 * 16}{\text{Time range} * \text{Phys CPU adjustment factor}}$$

If ActualMSU >= MSUatWeight-5% Then LPAR is actually capped

RMF z/OS 1.11 Enhancements for Group Capacity...

PARTITION DATA REPORT

z/OS V1R11

SYSTEM ID TRX1
RPT VERSION V1R11 RMF

DATE 02/26/2009
TIME 11.00.00

INTERVAL 05.00.000
CYCLE 1.000 SECONDS

MVS PARTITION NAME	TRX1	NUMBER OF PHYSICAL PROCESSORS	24	GROUP NAME	RMFGRP
IMAGE CAPACITY	60	CP	3	LIMIT	60
NUMBER OF CONFIGURED PARTITIONS	58	AAP	1	AVAILABLE	43
WAIT COMPLETION	NO	IFL	18		
DISPATCH INTERVAL	DYNAMIC	ICF	1		
		IIP	1		

----- PARTITION DATA -----							-- LOGICAL PARTITION PROCESSOR DATA --				-- AVERAGE PROCESSOR UTILIZATION PERCENTAGES --				
NAME	S	----MSU----			-CAPPING--		PROCESSOR- NUM	TYPE	----DISPATCH TIME DATA----		LOGICAL PROCESSORS		--- PHYSICAL PROCESSORS ---		
		WGT	DEF	ACT	DEF	WLM%			EFFECTIVE	TOTAL	EFFECTIVE	TOTAL	LPAR MGMT	EFFECTIVE	TOTAL
TRX1	A	400	100	4	NO	0.0	3.0	CP	00.00.11.049	00.00.11.371	1.23	1.26	0.04	1.23	1.26
H05LP45	A	10	0	2	NO	0.0	2	CP	00.00.04.720	00.00.05.690	0.79	0.95	0.11	0.52	0.63
TRX2CFA	A	100	0	1	YES	0.0	1	CP	00.00.02.958	00.00.03.078	0.99	1.03	0.01	0.33	0.34
H05LP59	A	100	0	1	NO	0.0	3	CP	00.00.02.700	00.00.03.501	0.30	0.39	0.09	0.30	0.39
H05LP60	A	10	0	9	NO	0.0	2	CP	00.00.23.742	00.00.26.331	3.96	4.39	0.29	2.64	2.93
TRX2	A	200	100	13	NO	0.0	3.0	CP	00.00.37.219	00.00.37.721	4.14	4.19	0.06	4.14	4.19
PHYSICAL									00.00.23.659				2.63		2.63

Field Heading	Meaning
AVAILABLE	Long-term average of CPU service units which would be allowed by the limit of the capacity group but are not used by its members. If the value is negative, this capacity group is subject to capping.

RMF z/OS 1.11 Enhancements for Group Capacity...

RMF Data Portal - Mozilla Firefox: IBM Edition

http://boetrx2.boeblingen.de.ibm.com:8803/

RMF Monitor III Data Portal for z/OS

RMF Report [,TRX2,MVS_IMAGE] : CPC (Central Processor Complex)

Time Range: 03/18/2009 08:46:00 - 03/18/2009 08:47:00

Partition Name: TRX2	CPU Type: 2097	CPU Model: 704	CPC Capacity (MSU/h): 401
Weight % of Max: 19.9	4h MSU Average: 2	Capacity Group Name: RMFGRP	Image Capacity: 60
WLM Capping %: 0.0	4h MSU Maximum: 3	Capacity Group Limit: 150	Less than 4h in Capacity Group: N
Proj Time until Capping: 14400	Proj Time until Group Capping: 14400	4h Unused Group Capacity Average: 142	CPC sequence number: 000000000001EBAE
# CP Processors: 4	# ICF+IFL+AAP Processors: 0	# AAP Processors: 1	# ICF Processors: 2
# IFL Processors: 18	# IIP processors: 1	Configured Partitions: 58	Wait Completion: NO
% Capacity Used: 7	# Dedicated CPs: 0	# Dedicated AAPs: 0	# Dedicated IIPs: 0
# Shared physical CPs: 4	# Shared physical AAPs: 1	# Shared physical IIPs: 1	Vary CPU management available: NO
WLM LPAR management enabled: YES	Physical Total % of shared CPs: 5.1	Physical Total % of shared AAPs: 0.0	Physical Total % of shared IIPs: 0.0
Physical Total % of shared ICFs: 61.1	Physical Total % of shared IFLs: 0.0		

Monitor III CPC report in Monitor III Data Portal displays the projected remaining time until image/group capping in the report header

Average available capacity for the group during last 4 hours

LPAR Name	Defined MSU/h	Actual MSU/h	Capping Option	# Logical Processors Online	Logical Effective %	Logical Total %	LPAR Mgmt %	Physical Effective %	Physical Total %	Line Type	# Online Processors Shared	# Online Processors Dedicated	Current LPAR Weight	Logical Processor Share %	Hiper Dispatch: # High	Hiper Dispatch: # Medium	Hiper Dispatch: # Low	Operating System Name	LPAR Cluster Name	Initial Weight	Mir We
*CP				14.0			2.5	4.8	7.3	CS	14	0	820								
H05LP45	0	1	NO	2.0	0.4	0.5	0.1	0.2	0.3	CP	2	0	10	2.4	N/A	N/A	N/A	BOEH0545			
H05LP59	0	1	NO	3.0	0.3	0.4	0.1	0.2	0.3	CP	3	0	100	16.2	N/A	N/A	N/A	BOEH0559			
H05LP60	0	8	NO	2.0	3.8	4.1	0.2	1.9	2.1	CP	2	0	10	2.4	N/A	N/A	N/A	BOEH0560			
TRX1	50	4	NO	3.0	1.2	1.3	0.0	0.9	1.0	CP	3	0	415	51.2	1	2	0	TRX1	TRX1PLEX	400	9
TRX2	60	4	NO	3.0	1.3	1.4	0.0	1.0	1.0	CP	3	0	185	90.2	0	1	2	TRX2	TRX1PLEX	200	9
TRX2CFA	0	2	YES	1.0	2.1	2.1	0.0	0.5	0.5	CP	1	0	100	48.7	N/A	N/A	N/A				
PHYSICAL									2.2	CY											
*ICFPOOL				1.0			1.4	56.8	58.2	IS	198	1	0								

RMF z/OS 1.11 Enhancements for Group Capacity...

New DDS metrics

1EBAE,TRX2,LPAR -- available capacity (MSU/h) for group [8D43E0]
 Time Range: 04/16/2009 14:01:00 - 04/16/2009 14:02:00
 9

1EBAE,TRX2,LPAR -- remaining time until group capping in seconds [8D4460]
 Time Range: 04/16/2009 14:01:00 - 04/16/2009 14:02:00
 14400

This window will automatically refresh every 60 seconds (MINTIME) ...

Automatic refresh in 8 seconds (MINTIME) ... Local intranet

remaining time until group capping in seconds (by partition)

The projected time until WLM considers to cap members of the capacity group. That is, the usage of processor resources for one or more members of the group might be limited. WLM soft capping takes place to prevent you from using more than the defined group capacity limit over a long period of time. This is under the assumption you continue to use your system as you have done in the immediate past. The maximum number RMF reports is 14400 seconds or 4 hours. If RMF reports 14400, it means the remaining time until the group becomes subject to capping is at least 4 hours.

available capacity (MSU/h) for group (by partition)

The long-term average of CPU service in MSUs/h which would be allowed by the limit of the capacity group but is not used by its members. If the value is zero or negative, WLM considers to cap members of the group.

Group Capacity: Availability

Function	z/OS V1.12 as previewed 2/2010	z/OS V1.11	z/OS V1.10	Earlier Releases
Group Capacity plus OA24096 Enhancements	+	+	OA24096 OA23230	OA24096 OA23230 (z/OS 1.8)
RMF Reporting Enhancements for Group Capacity	+	+		
z/OS Capacity Provisioning	+	+	OA20824	

- **OA24096**
 - Changes the behavior when then group limit is changed according to the behavior for an individual defined capacity limit
- **OA23230**
 - Corrects a storage overlay which will occurs when SMF 99 data is collected and a partition is dynamically activated via HCD
- **Short Comings of the existing Group Capacity Report**
 - Reporting was not sufficient to understand capping of partitions within a group
 - Resolved with z/OS 1.8 RMF Reporting Enhancements
- **Related z/OS Functions**
 - z/OS Capacity Provisioning allows to activate additional CPU capacity via OOCoD in a controlled manner.

WLM Reporting: Extend Number of Report Classes



- Problem encountered

WLM supported at most 999 report classes which has become insufficient for large installations.

- Solution

Extend number of report classes in multiple steps:

First Step (z/OS 1.11):

Extend to 2047 Report Classes.

Expand internal data structures to be able to deal with 4095 report classes.

Second Step:

Extend to 4095 (the maximum possible value) Report Classes in future release.

Why do we need multiple steps ??

This is to avoid compatibility issues when running a sysplex with lower level releases (z/OS 1.10 and earlier cannot properly handle more than 2047 report classes).

- Annotation

New WLM functionality level in z/OS 1.11: LEVEL023

For Service Definitions in XML format, the corresponding XML namespace is
<http://www.ibm.com/xmlns/prod/zwlm/2009/09/ServiceDefinition.xsd>

Extended Number of Report Classes

Availability

Function	z/OS V1.12	z/OS V1.11	z/OS V1.10	Older Releases
2047 Report Classes	+	+		

New Programming Interface for Monitors

Control Block: IRARMCTZ

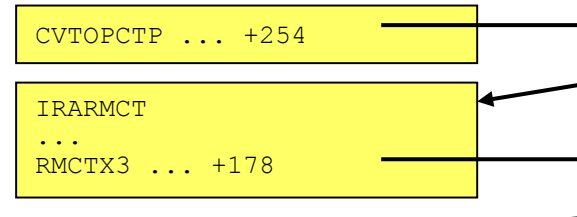
- New extension to SRM Control Table (PI) for information which is of interest for externalization
 - For example all information related to RMF's Monitor II OPT report is included in this table

```

boewlm1 - wc3270
RMF - OPT Settings                               Line 1 of 29
Command ==>                                     Scroll ==> PAGE
CPU= 4      UIC= 65K PR= 0                      System= WLM1 Total

OPT: 00      Time: N/A
-- Parameter -- Default -- Value -- Unit ----- Description -----
ABNORMALTERM      Yes      Yes Y/N  Abnormal terminations in routing
BLWLINTHD         20      20 sec  Time blocked work waits for help
BLWLTRPCT         5        5 0/00  CPU cap. to promote blocked work
CCCAWMT          12000    12000 usec  Alternate wait management time
ZAAPAWMT          12000    12000 usec  AWM time value for ZAAPs
ZIIPAWMT          12000    12000 usec  AWM time value for ZIIPs
CNTCLIST          No       No Y/N  Clist commands count individually
CPENABLE         10,30|0,0  10,30 %  Threshold for TPI (low,high)
DVIO              Yes      Yes Y/N  Directed VIO is active
ERV              500      500/CB SU  Enqueue residency CPU Service/DP
HIPERDISPATCH    No       No/No Y/N  Hiperdispatch is desired/active
IFAHONORPRIORITY  Yes      Yes Y/N  Allows CPs to help ZAAPs
IIPHONORPRIORITY  Yes      Yes Y/N  Allows CPs to help ZIIPs
INITIMP           0        0/FE #   INITIMP value/DP for initiators
IRA405I           70,50,50  70,50,50 %  Fixed storage of <16M,16M-2G,tot
WAXPROMOTETIME   6         6 *10s   Holder allowed to run promoted
WCCAFCTH         400,800   400,800 #  Threshold for storage (low,ok)
WCCFXEPR         92        92 %     Fixed storage threshold < 16 MB
WCCFXTPR         80        80 %     Fixed online storage threshold
PROJECTCPU       No       No Y/N  CPU projection for ZAAPs, zIIPs
RCCFXET         82,88     82,88 %  Physical MPL threshold (low,high)
RCCFXIT         66,72     66,72 %  Logical MPL threshold (low,high)
RMPTTOM         1000|3000 3000 msec  SRM invocation interval
RTPIFACTOR       100       100 %    PI affects server routing weights
STORAGENSWDP     Yes      Yes Y/N  Sets non-swap. ASID non-dispatch.
STORAGEWTOR     Yes      Yes Y/N  WTOR to cancel AS in shortage
VARYCPU         Yes      Yes Y/N  VARYCPU is enabled
VARYCPUMIN       1        1 #      VARYCPUMIN value
WASROUTINGLEVEL  0         0 #      WebSphere routing level

F1=HELP      F2=SPLIT  F3=END    F4=RETURN  F5=RFIND   F6=SORT
F7=UP        F8=DOWN   F9=SWAP   F10=LEFT   F11=RIGHT  F12=RETRIEVE
4B          X
IPY$1C09    002/015
    
```



IRARMCTZ Map

Offsets		Type/Value	Len	Name (Dim)	Description
Dec	Hex				
0	(0)	STRUCTURE	0	RMCTZ	
0	(0)	CHARACTER	8	RMCTZ_NAME	control block acronym > IRARMCTZ <
8	(8)	BITSTRING	1	RMCTZ_VERSION	
9	(9)	BITSTRING	1		Rmctz version Reserved
10	(A)	SIGNED	2	RMCTZ_LENGTH	Size of RMCTZ
12	(C)	BITSTRING	1	RMCTZ_LPAR_FLAGS	
					(0)
					LPAR Management flags updated by SRM.
					RMCTZ_LPARMGMT_ENABLED
					"X80" ON if WLM LPAR Management Processing is enabled
					RMCTZ_LPAR_VARYCPU_ENABLED
					"X40" ON if VARYCPU option is turned on either by default or is explicitly set to 'on'
13	(D)	CHARACTER	1	RMCTZ_FLAG1	RMCTZ Flag 1
					(0)
					RMCTZ_ABN_OPT
					"X80" ABNORMALTERM option set
					RMCTZ_FLAG1_RSVD1
					"X7F" reserved
14	(E)	CHARACTER	1	RMCTZ_FLAG2	RMCTZ Flag 2
					(0)
					RMCTZ_VCM_OPT
					"X80" 1:=VCM specified
					RMCTZ_VCM
					"X40" 1:=Running in vertical CP management mode
					RMCTZ_FLAG2_RSVD1
					"X3F" reserved

New Programming Interface for Monitors: Availability

Control Block: IRARMCTZ

Function	z/OS V1.12 as previewed 2/2010	z/OS V1.11	z/OS V1.10	Earlier Releases
RMF Monitor II OPT Display	+	+		
WLMOPT Tool (bundled with WLMQUE Tool)	No longer extended Still bundled with WLMQUE but on z/OS 1.10 level		+	Since z/OS 1.8
IRARMCTZ	+	OA31201	OA31201	

- **RMF Monitor II OPT Display**
 - Replaces WLMOPT Tool
 - Bundled with WLMQUE Tool but no longer extended (remains on z/OS 1.10 level)
 - WLMQUE Tool is still valid (see also WLM Tools summary)

- **New data interface for Monitors**
 - Introduced with z/OS 1.12
 - Rollback to z/OS 1.10

Enhanced Storage Monitoring

■ Problem

- Pageable and Auxiliary storage shortages can lead to serious system problems (including system outages).

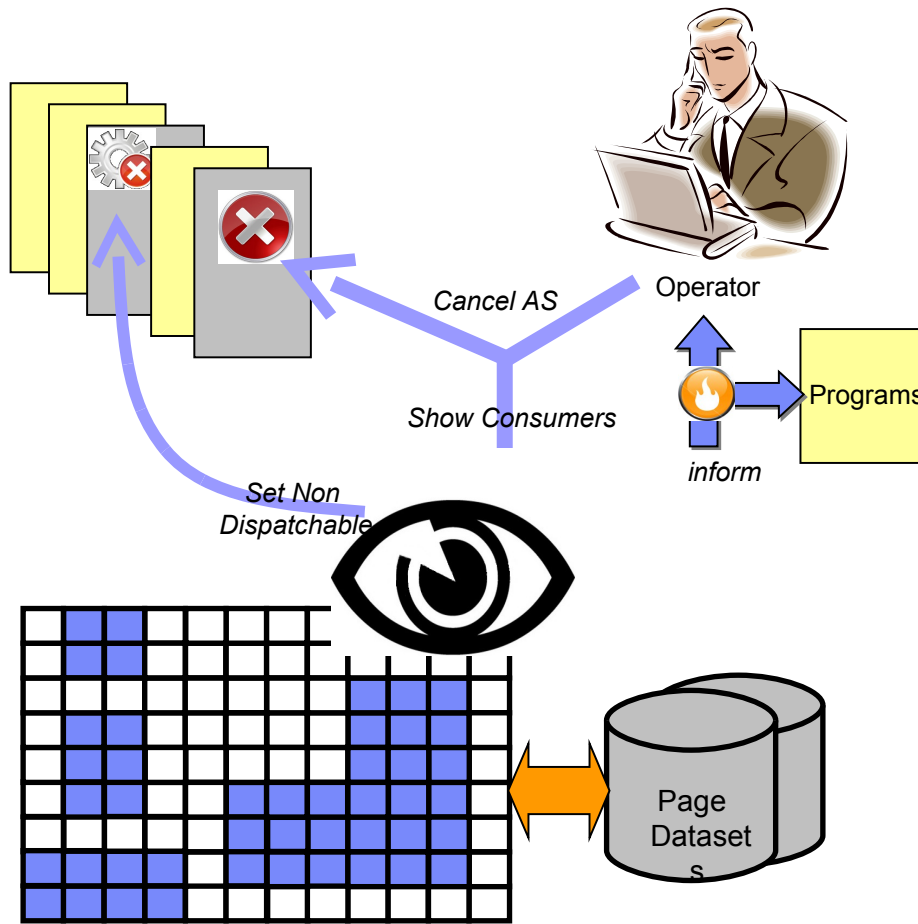
■ What needs to be done

- Identify a storage shortage when it occurs.
- Identify the reason (causing application) of the storage shortage.
- Give the installation a chance to react on a storage shortage when it occurs.

■ Solution

- Introduce a new set of messages to warn the installation when auxiliary and pageable storage shortages occur.
- Introduce a WTOR which allows the installation to cancel storage consumers.
- Set storage consumers non-dispatchable to allow the installation to react on the situation.
- Introduce a set of new programming interfaces (ENF signal) to allow applications to react on storage shortages.

Storage Shortage Management



Monitors

- Fixed Storage consumption
- Auxiliary Storage consumption
- Every 2 seconds

Informs in case of problems

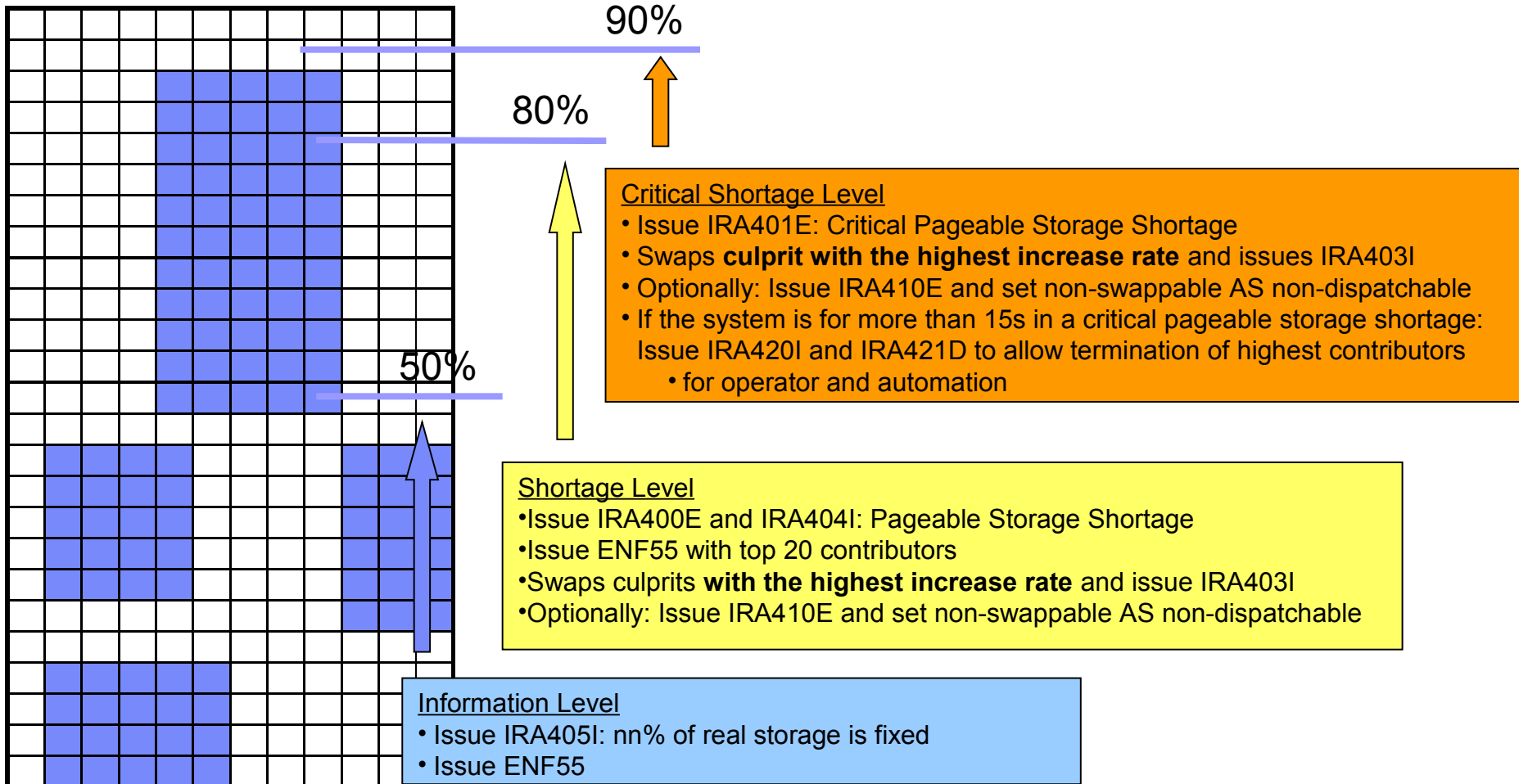
- Operator via messages
- Programs via ENF55

Takes Actions

- To set Address Spaces non dispatchable
- To cancel address spaces on operator request

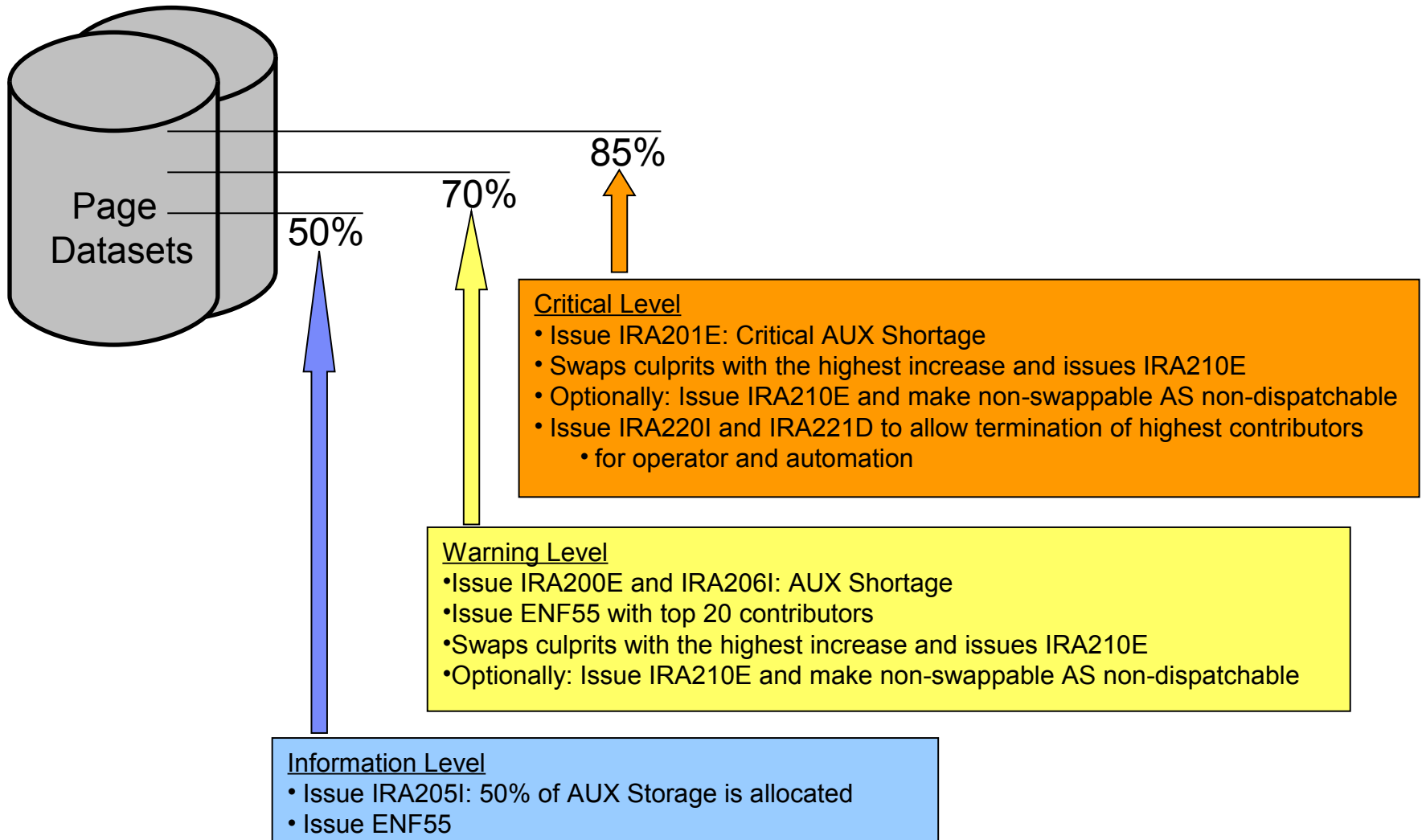
Pageable Storage Shortages – Details ...

Real Storage:



Note: for Below 16M the shortage targets are 92% and 96%

Auxiliary Storage Shortages – Warning Levels



Storage Enhancements: Availability

Function	z/OS V1.12	z/OS V1.11	z/OS V1.10	Earlier Releases
Reserve Frames for CHNGDUMP command	+	+		
Enhanced Storage Monitoring	+	+	+	
Support for >128GB Real Storage	+	+	+	z/OS 1.8

- Enhanced Storage Monitoring is introduced for z/OS 1.10
- Support for >128GB Real Storage was introduced with z/OS 1.8
- Example: `CHNGDUMP SET,SDUMP,BUFFERS=1K`
 - Reserves frames on the available frame queue for Dump processing (in the example 256 frames)

WLM Tools: A Summary

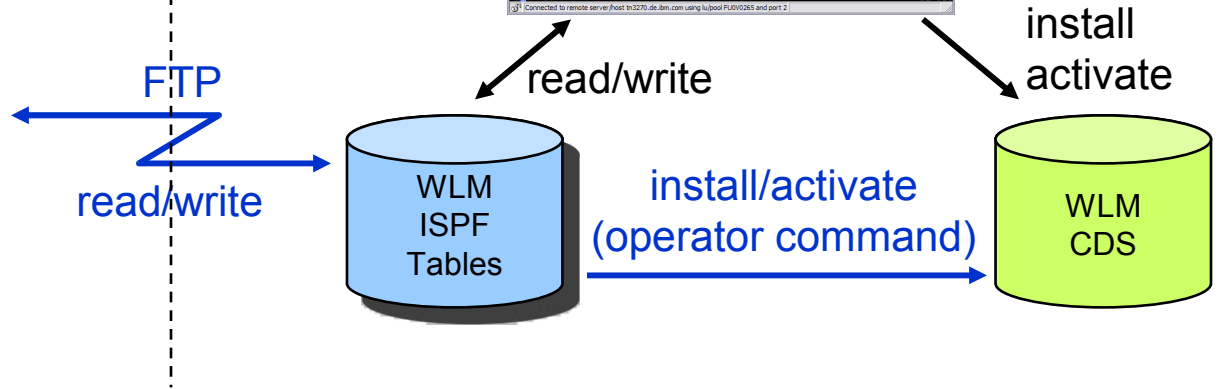
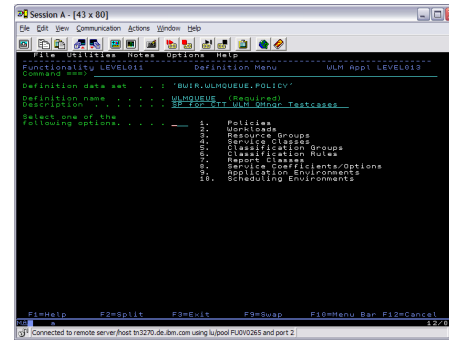
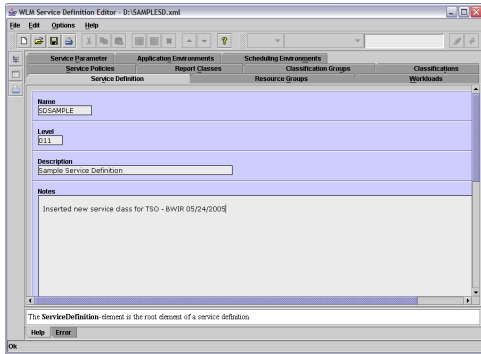
Tool	Name	Description	Content	Support
MIGRATE	Goal Mode Migration Aid	Assists migration from compatibility to goal mode	Excel/workstation SMF/RMF processing/MVS	No!! Removed from WLM Tools page in 2004
WLMZOS	OS/390 to z/OS Execution Velocity Migration	With z/OS the using samples are no longer sampled but calculated. The tool helps to understand whether this changes the amount of delays for service classes with execution velocity goals	Excel/workstation SMF processing tool on MVS	No!! Removed from WLM Tools page in 2006
SVDEF	Service Definition Formatter	Uses output from WLM Administrative Administration to display content of service definition in a workstation spreadsheet	Excel/workstation	Not updated anymore but still available on WLM Tools page
WSE	Service Definition Editor	Allows to create, mmodify, retrieve and install WLM service definitions	Java program on workstation	YES!! Available
WLMQUE	Application Environment Viewer	Allows to monitor WLM Application Environments	ISPF Tool	YES!! Available
WLMOPT	OPT Display	Display WLM/SRM OPT Parameters	IPF Tool	No!! Replaced with z/OS 1.11 by RMF

<http://www-03.ibm.com/servers/eserver/zseries/zos/wlm/tools/>

WLM Tools

Service Definition Editor

Workstation z/OS



WLM Tools

Service Definition Editor

WLM Service Definition Editor - D:\SAMPLESD.xml

File Edit Options Help

local

Classification Groups			Classifications		Service Parameter		Application Environments			Scheduling Environments	
Service Definition			Resource Groups		Workloads		Service Policies			Report Classes	
Name	ServiceClasses	Period	Goal	Im...	Duration	ResponseTime	Perce...	Level	ResourceGr...	CPU...	Description
WKLDASC											ALL APPC Transaction
WKLDASC	A3V30STD								GBATCH20	No	ASCH default Service Class
WKLDASC	A3V30STD	1	Velocity	2	500			10			
WKLDASC	A3V30STD	2	Velocity	2	-			8			
WKLTIK											All Batch Jobs
WKLTIK	B4V10STD								-	No	Batch Standard VEL 10 IMP 4
WKLTIK	B4V10STD		velocity	4	-			10			
WKLTIK	B4V20STD		velocity	4	-			10		No	Batch Standard VEL 20 IMP 4
WKLTIK	B4V20STD	1	Velocity	4	1000			20			
WKLTIK	B4V20STD	2	Velocity		-			10			
WKLDTSO											ALL TSO USERIDS
WKLDTSO	T2335DEV								-	No	Developer (Standard) TSO
WKLDTSO	T23		PrecentileResponseTime	2	2500	00:00:02.000	98				
WKLDTSO	T23		AverageResponseTime	3	300000	00:00:20.000	95				
WKLDTSO	T23		PercentileResponseTime	5	-			10			
WKLDTSO	T23								-	No	Production TSO Helpers
WKLDTSO	T23		PrecentileResponseTime	2	2000	00:00:01.000	99				
WKLDTSO	T23		PrecentileResponseTime	2	10000	00:00:02.000	99				

used by:
Classification: JES

- Insert
- Insert Before
- Insert After
- Replace by
- Copy
- Cut
- Delete

No	Description	Element
1	Importance value can not be null	Workload "WKLTIK"/ServiceClass "B4V20STD"/Velocity (#1)
2	WLM may not distinguish between periods with equal importance and only slightly different velocity levels	Workload "WKLDASC"/ServiceClass "A3V30STD"

Help Error

Ok

WLM Tools

Display WLM/SRM OPT Parameter (WLM Tool, supported up to R10)

```

Command ==>                                Scroll ==> PAGE
.
.
.      WLM OPT Settings                                >SAVE<
System: AQFT      Version: z/OS 011100  OPT: FT  Time: not issued
OPT-Parameter:   Value:                Description:
.
.
ABNORMALTERM      Yes  Abnormal term. used in routing rec.
BLWLTRPCT         5    CPU cap. to promote blocked work
BLWLINTHD         20   Time blocked work waits for help
CCCAWMT           3200,3200  AWM time value (defined, used)
ZAAPAWMT          3200,3200  AWM time value for zAAPs (def, used)
ZIIPAWMT          3200,3200  AWM time value for zIIPs (def, used)
CNTCLIST          No    Clist commands count individually
CPENABLE          10,30  LOW,HI thresh for % TPI int. x 100
DVIO              Yes   Specifies w/ directed VIO is active
ERV              1000,E6  Eng res. CPU Service and DP
HIPERDISPATCH   Yes,Yes Hiperdispatch value(inOPT, Running)
IFAHONORPRIORITY Yes   Specifies if CPs may help zAAPs
IIPHONORPRIORITY Yes   Specifies if CPs may help zIIPs
INITIMP          0,FE   INITIMP value and DP for initiators
MCCAFCTH         400,800  LOW,HIGH central threshold
MCCFXEPR         92     % of storage fixed within first 16MB
MCCFXTPR         80     % of online storage fixed
PROJECTCPU       No    CPU projection for zAAPs and zIIPs
RCCFXTT          66,72  Low,High Logical MPL threshold
RCCFXET          82,88  Low,High Physical MPL threshold
RMPTTOM          1000   SRM invocation interval
STORAGESENSDP    Yes   Set Non-swappable AS non dispatchable
STORAGEWTOR      Yes   Issue IRA221D and IRA421D
IRA405I          46,32,32  IRA405I warning level: 16M,2G,Tot
VARYCPU          No    VARYCPU is enabled
VARYCPUMIN       1     VARYCPUMIN value
WASROUTINGLEVEL  0     WebSphere Routing Level
.
.

```

WLM Tools

Display WLM/SRM OPT Parameter (RMF Monitor II OPT Report)

```

boewlm1 - wc3270
RMF - OPT Settings
Line 1 of 29
Command ==>
Scroll ==> PAGE
CPU= 4    UIC= 65K PR= 0    System= WLM1 Total

OPT: 00    Time: N/A
-- Parameter -- - Default - -- Value -- Unit ----- Description -----
ABNORMALTERM      Yes      Yes Y/N  Abnormal terminations in routing
BLWLINTHD         20      20 sec  Time blocked work waits for help
BLWLTRPCT         5        5 0/00  CPU cap. to promote blocked work
CCCAWMT          12000   12000 usec  Alternate wait management time
ZAAPAWMT         12000   12000 usec  AWM time value for ZAAPs
ZIIPAWMT         12000   12000 usec  AWM time value for zIIPs
CNTCLIST          No       No Y/N  Clist commands count individually
CPENABLE         10,30|0,0 10,30 %  Threshold for TPI (low,high)
DVIO              Yes      Yes Y/N  Directed VIO is active
ERV              500     500/CB SU  Enqueue residency CPU Service/DP
HIPERDISPATCH   No       No/No Y/N  Hiperdispatch is desired/active
IFAHONORPRIORITY Yes      Yes Y/N  Allows CPs to help zAAPs
IIPHONORPRIORITY Yes      Yes Y/N  Allows CPs to help zIIPs
INITIMP          0        0/FE #   INITIMP value/DP for initiators
IRA405I          70,50,50 70,50,50 %  Fixed storage of <16M,16M-2G,tot
VAXPROMOTETIME   6        6 *10s  Holder allowed to run promoted
WCCAFCTH         400,800 400,800 #   Threshold for storage (low,ok)
WCCFXEPR         92       92 %     Fixed storage threshold < 16 MB
WCCFXTPR         80       80 %     Fixed online storage threshold
PROJECTCPU       No       No Y/N  CPU projection for zAAPs, zIIPs
RCCFXET         82,88   82,88 %  Physical MPL threshold (low,high)
RCCFXTT         66,72   66,72 %  Logical MPL threshold (low,high)
RMPTTOM         1000|3000 3000 msec  SRM invocation interval
RTPIFACTOR       100     100 %    PI affects server routing weights
STORAGENSWDP     Yes      Yes Y/N  Sets non-swap. ASID non-dispatch.
STORAGEWTOR     Yes      Yes Y/N  WTOR to cancel AS in shortage
VARYCPU          Yes      Yes Y/N  VARYCPU is enabled
VARYCPUMIN       1        1 #     VARYCPUMIN value
WASROUTINGLEVEL  0        0 #     WebSphere routing level

F1=HELP      F2=SPLIT      F3=END      F4=RETURN      F5=RFIND      F6=SORT
F7=UP        F8=DOWN       F9=SWAP lis F10=LEFT      F11=RIGHT     F12=RETRIEVE
4B X        T        IPY$1C09    002/015
  
```

WLM Tools

WLMOPT – WLM Application Environment Viewer

```

Command ==>                               Scroll ==> PAGE
Application Environment Monitor
Selection: >HELP< >SAVE< >OVW< >ALL< \AE=SYSBATCH
System: AQFT      Sysplex: MCLXCF01  Version: z/OS 011100  Time: 06:22:27

ApplEnv_ Type SubName_ WMAS Del Dyn NQ QLen Str Hav Unb Trm Min_ Max_ ICnt
SYSBATCH JES  JES2      0031 No No  3   0  0 12  0  0  0  0  0

WorkQue_ Del Wnt Hav ICnt QueIn_ QueOut QueLen QueTot_ Act_ Idl_
WLMLONG  No  7  7  0  0  0  0  0  0  4  3
WLMSHORT No  3  3  0  0  0  0  0  0  2  0
COMBUILD No  2  2  0  0  0  0  0  0  1  1

SvAS Binding_ Ter Opr Btc Dem Have Jobname
0043 WLMLONG  No No Yes No 1 BCNDEVD
0175 WLMLONG  No No Yes No 1 ALLAEBS.2.SEAS.2.JBNI
0166 WLMLONG  No No Yes No 1 SERV9956
0165 WLMLONG  No No Yes No 1 SERV9955
015A COMBUILD No No Yes No 1 C90SPACE
0150 WLMLONG  No No Yes No 1 INIT
0202 WLMLONG  No No Yes No 1 INIT
0152 COMBUILD No No Yes No 1 INIT
0229 WLMSHORT No No Yes No 1 BMGX1$
0119 WLMLONG  No No Yes No 1 INIT
0050 WLMSHORT No No Yes No 1 ALLAEBS.2.SEAS.11.JBNI
01A5 WLMSHORT No No Yes No 1 INIT

```

Contact Information

Thomas Blattmann

z/OS WLM/SRM Development

IBM Deutschland Research & Development

71032 Böblingen, Germany

E-Mail: blattmann@de.ibm.com

धन्यवाद

Hindi

多謝

Traditional Chinese

ขอบพระคุณ

Thai

Спасибо

Russian

Gracias

Spanish

Thank You

English

Obrigado

Brazilian Portuguese

شكراً

Arabic

多谢

Simplified Chinese

Danke
German

Bedankt

Dutch

Grazie

Italian

Merci
French

நன்றி

Tamil

ありがとうございました

Japanese

감사합니다

Korean