SHARE
Technology · Connections · Results

z/OS Workload Management (WLM) Update for z/OS V2.1 and V1.13

Horst Sinram
IBM Germany Research & Development
z/OS Workload Management

August, 2013

Session 14211

SHARE • in Boston

Trademarks

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

AIX*	DS8000*	Language Environment*	SystemPac*	z10
BladeCenter*	FICON*	Parallel Sysplex*	System Storage	z10 BC
DataPower*	HiperSockets	POWER7*	System z	z10 EC
DB2*	Hyperwap	PrintWay	System z9	z/OS*
DFSMS	IBM*	ProductPac*	System z10	zEnterprise
DFSMSdss	IBM eServer	RACF*	System z10 Business Class	zSeries*
DFSMShsm	IBM logo*	REXX	WebSphere*	
DFSMSrmm	ibm.com	RMF	z9*	
DFSORT	Infiniband*	ServerPac*		
DS6000*	InfoPrint			

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

InfiniBand is a registered trademark of the InfiniBand Trade Association (IBTA).
Intel is a trademark of the Intel Corporation in the United States and other countries.
Linux is a trademark of Linux Torvalds in the United States, other countries, or both.
Java and all Java-related trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc., in the United States and other countries.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
UNIX is a registered trademark of The Open Group in the United States and other countries.
All other products may be trademarks or registered trademarks of their respective companies.
The Open Group is a registered trademark of The Open Group in the US and other countries.

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.

This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.

Agenda

- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- I/O Priority Groups
- Other z/OS V2.1 Enhancements
- Manage CICS Regions Using Goals Of: "BOTH"
- Response Time Distribution for Velocity Goals
- HiperDispatch Enhancements

IBM zEnterprise EC12 GA2 Support Overview

- zEnterprise EC12 (zEC12) offers new functions for hard and soft capping with GA2 (firmware driver 15):
  - Smoother capping with WLM managed softcapping
  - When IRD weight management is active the group capacity of an LPAR may be derived by the initial weight
  - New "Absolute Capping Limit" LPAR control

| z/OS release / Function	V2.1	V1.13	V1.12
Smoother capping	**+**		
Group capacity to use initial weight	**+**	OA41125	OA41125
Absolute capping	**+**	OA41125	OA41125

Capping algorithms today

Pattern capping

- Must be used when $MSU@LPARweight <$ definedLimit
- Periods with LPAR capped at weight and running uncapped
- Can result in "pulsing" potentially impacting online workloads

MSU@Weight < Limit
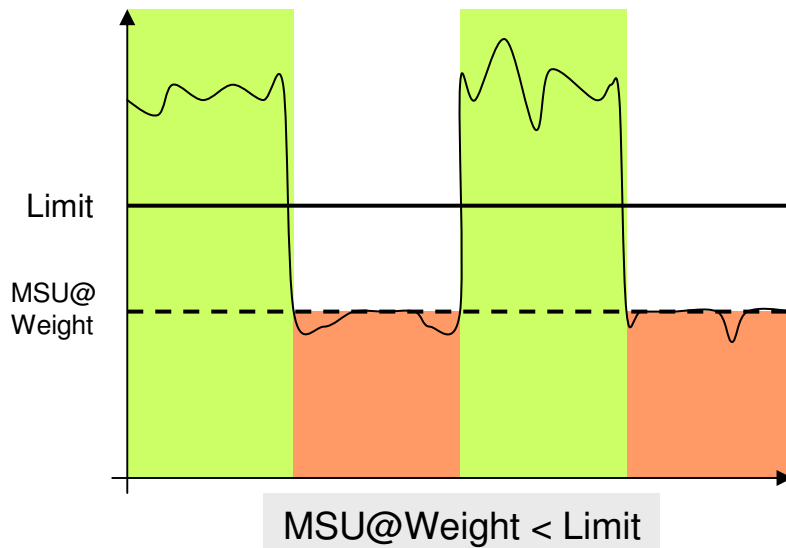
Phantom weight capping

- Is used when $MSU@LPARweight \geq$ definedLimit
- Internally PR/SM uses an additional weight to limit LPAR consumption below weight
  - Phantom weight must be non-negative pre-zEC12 GA2
- Results in smooth capping

MSU@Weight > Limit
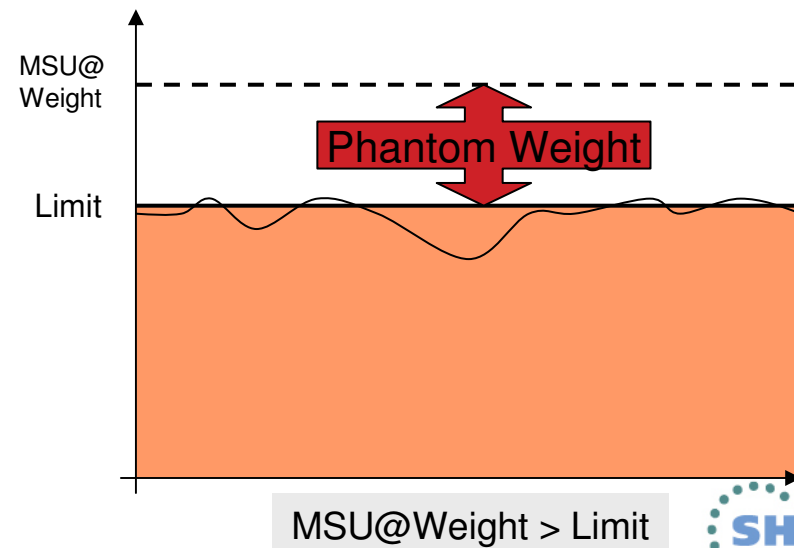
zEC12 GA2 Negative Phantom Weight

- zEC12 GA2 allows using a negative phantom weight for soft capping

- Therefore, when *MSU@LPARweight* < definedLimit WLM can now use a negative phantom weight instead of pattern capping
  - **I.e.,** phantom weight capping becomes the only mechanism

- z/OS V2.1 will exploit this feature
  - Eliminates pulsing effects caused by cap patterns

zEC12 GA2 can use initial weight for group capping

- It is possible to combine IRD weight management with capacity groups
  - IRD changes the –current- weight in order to shift capacity within an LPAR cluster

- However, IRD weight management gets suspended when capping is in effect
  - Because entitlement of an LPAR within a capacity group is currently derived from the current weight the LPAR might get stuck at a low weight
    - Consequently, a low group capacity entitlement can result

- On zEC12 GA2 the **initial** LPAR weight will be used for group capacity
  - Only if all systems in a capacity group are
    - z/OS V2.1, or
    - z/OS V1.12, V1.13 with OA41125 applied.
  - Results in more predictive and better controllable group capacity entitlement

zEC12 GA2 Absolute Capping Limit

- zEC12 GA2 allows specification of an "absolute capping limit"
  - Primarily intended for non z/OS images
  - Expressed in terms of 1/100ths of a processor
  - Can be specified independently from the LPAR weight
    - But recommended to specify absolute cap above weight
  - Therefore, it is insensitive to LPAR (de)activations and less sensitive to capacity changes
  - Can be specified per processor type in image profile and partition controls panel

- Unlike initial capping absolute capping may be used *concurrently* with defined capacity and/or group capacity management
  - The respective minimum becomes effective.
  - WLM/SRM recognizes new cap, e.g. for routing decisions.
  - RCTIMGWU = MIN(absolute cap, defined capacity, group cap) when all capping types are in effect
    - RMF provides RCTIMGWU in SMF70WLA
    - In addition, SMF70HW_Cap_Limit value in hundredths of CPUs

zEC12 GA2 Absolute Capping Limit - Examples

Change Logical Partition Controls - P35

Last reset profile attempted:
Input/output configuration data set (IOCDS):A0 198AP35

Tabs: **CPs** | zAAPs | IFLs | zIIPs | Processor Running Time

Logical Partitions with Central Processors

--- Select Action ---

Logical Partition	Active	Defined Capacity	WLM	Current Weight	Initial Weight	Min Weight	Max Weight	Current Capping	Initial Capping	Absolute Capping	Number of Dedicated Processors	Number of Not dedicated Processors
IRD6	Yes	10	☐	300	300			No	☑	3.20	0	3

Customize Image Profiles: IRD8

- IRD8
  - IRD8
    - General
    - Processor
    - Security
    - Storage
    - Options
    - Load
    - Crypto

Group Name

Logical Processor Assignments

☐ Dedicated processors

Select	Processor Type	Initial	Reserved
☑	Central processors (CPs)	3	1
☑	System z application assist processors (zAAPs)	0	1
☑	System z integrated information processors (zIIPs)	0	1

Not Dedicated Processor Details for :

◉ CPs ○ zAAPs ○ zIIPs

CP Details

Initial processing weight   80   1 to 999   ☐ Initial capping

☑ Enable workload manager

   Minimum processing weight   30

   Maximum processing weight 700

Absolute Capping   ○ None

       ◉ Number of processors (0.01 to 255.0) 2

Agenda

- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- 3000 Application Environments
- I/O Priority Groups
- Other z/OS V2.1 Enhancements
- Manage CICS Regions Using Goals Of: "BOTH"
- Response Time Distribution for Velocity Goals
- HiperDispatch Enhancements

New Classification Qualifiers and Groups: Overview

- With z/OS V2R1, WLM/SRM introduces
  - New classification group types, and
  - Some new and modified work qualifier types for use in classification rules in the WLM service definition

- These can be used to improve readability of the WLM service definition for work when there is no naming convention that allows masking or wild-carding

- New and modified qualifier types allow better classification of new DB2 and DDF workload

- More notepad information about a service definition allowed

New Classification Qualifiers and Groups

- z/OS V2.1 extends classification groups for all non-numeric work qualifier types.
- For long qualifier types, a start position for group members, and nesting is allowed.
- **New Groups:**

  - Accounting Information Group
  - Client Accounting Information Group
  - Client IP Address Group
  - Client Transaction Name Group
  - Client Userid Group
  - Client Workstation Name Group
  - Collection Name Group

  - Correlation Information Group
  - Procedure Name Group
  - Process Name Group
  - Scheduling Environment Group
  - Subsystem Collection Group
  - Subsystem Parameter Group
  - Sysplex Name Group

New Classification Qualifiers and Groups

- Subsystems (DB2) require longer and additional work qualifiers:
  - Work qualifier type "Package Name":   128 characters (instead of 8)
  - Work qualifier type "Procedure Name":   128 characters (instead of 18)

  - New work qualifier types:
    - Client Accounting Information   (max. 512 characters)
    - Client IP Address   (max.   39 characters)
    - Client Transaction Name   (max. 255 characters)
    - Client User ID   (max. 128 characters)
    - Client Workstation Name   (max. 255 characters)

- The maximum number of "Notepad" lines the has been increased from 500 to 1000 lines

- Note: New and modified work qualifier types are only supported by the new 64-bit classify IWM4CLSY (planned to be used by DB2 V11).

WLM ISPF application enhancements

- Option 5 Classification Groups: Groups can be defined for all non-numeric work qualifier types.
  - Exceptions: Priority (numeric), zEnterprise Service Class

```
 File   Utilities   Notes   Options   Help
-----------------------------------------------------------------------
Functionality LEVEL029          Definition Menu          WLM Appl LEVEL029
Command ===> _____

Definition data set  . . . : none

Definition name  . . . . . . coeffs    (Required)
Description  . . . . . . . . Service coefficients

Select one of the
following options. . . . . 5    1.    Policies
┌──────────────────────────────────────────────────────────────────────┐
│                     Classification Group Menu                          │
│                                                                        │
│  Select one of the following options.                                  │
│ _    1.  Accounting Information Groups     14. Plan Name Groups         │
│      2.  Client Accounting Info Groups     15. Procedure Name Groups    │
│      3.  Client IP Address Groups          16. Process Name Groups      │
│      4.  Client Transaction Name Groups    17. Scheduling Environment Groups │
│      5.  Client Userid Groups              18. Subsystem Collection Groups │
│      6.  Client Workstation Name Groups    19. Subsystem Instance Groups │
│      7.  Collection Name Groups            20. Subsystem Parameter Groups │
│      8.  Connection Type Groups            21. Sysplex Name Groups       │
│      9.  Correlation Information Groups     22. System Name Groups       │
│     10.  LU Name Groups                    23. Transaction Class Groups  │
│     11.  Net ID Groups                     24. Transaction Name Groups   │
│     12.  Package Name Groups               25. Userid Groups            │
│     13.  Perform Groups                                                 │
│  F1=Help          F2=Split         F5=KeysHelp      F9=Swap       F12=Cancel │
└──────────────────────────────────────────────────────────────────────┘
```

WLM ISPF application samples

SHARE
Technology · Connections · Results

```
  Group   Xref   Notes   Options   Help
-------------------------------------------------------
                        Modify  a  Group
Command ===>  _____

Enter  or  change  the  following  information:

Qualifier  type  .  .  .  .  .  .  .  : Accounti
Group  name  .  .  .  .  .  .  .  .  : SLOWACCT
Description  .  .  .  .  .  .  .  .  .  .
Fold  qualifier  names?  .  .  .  . Y  (Y  or

Qualifier  Name   Start   D
020175_____          _____
030275_____          _____
040375_____          _____
```

> Use to group work when there is no naming convention that allows for masking or wild-carding

```
  Group   Xref   Notes   Options   Help
-------------------------------------------------------
                        Modify  a  Group
Command ===>  _____

Enter  or  change  the  following  information:

Qualifier  type  .  .  .  .  .  .  .  : Accounting  Information
Group  name  .  .  .  .  .  .  .  .  : FASTDEPT
Description  .  .  .  .  .  .  .  .  .  .  ___
Fold  qualifier  names?  .  .  .  . Y  (Y  or  N

Qualifier  Name   Start   Description
PURCHASE____      8___    _____
SALES_____      8___    _____
SHIPPING____      8___    _____
ITDEP*_____      11__    _____
HRDEP*_____      11__    _____
```

> Use a start position for each group member to indicate how far to index into the character string for a match. The start position needs not be the same for all group members.

Use of New Groups and Qualifiers in the WLM Administrative Application

New work qualifier types:
- Client Accounting Information
- Client IP Address
- Client Transaction Name
- Client User ID
- Client Workstation Name

Groups of long work qualifier types can be nested

Increased maximum length for work qualifier types Package Name and Procedure Name.

Classification via new groups: Examples

```
          1   AIG        SLOWACCT  ___              SLOW          _____
          2     AIG        FASTDEPT ___             FAST          _____
          1   CAI        CLIENTAI  ___              VEL20         _____
          1   CIP        CLIENTIP  ___              VEL30         _____
          1   CTN        CLIENTTN  ___              VEL40         _____
          1   CUI        CLIENTUI  ___              VEL50         _____
```

```
Accounting Information Group SLOWACCT -

Created by user  IBMUSER on 2011/08/23 at
Last updated by user  IBMUSER on 2011/08/

  Qualifier    Starting
  name         position     Description
  ---------    --------     ---------------
  020175
  030275
  040375
```

```
Accounting Information Group FASTDEPT -

Created by user  IBMUSER on 2011/08/23 a
Last updated by user  IBMUSER on 2011/08,

  Qualifier    Starting
  name         position     Description
  ---------    --------     ---------------
  PURCHASE     8
  SALES        8
  SHIPPING     8
  ITDEP*       11
  HRDEP*       11
```

- '040375,SHIPPING' → FAST.

- '030275,D71ITDEP' → FAST.

- '020175,CONTROL ' → SLOW, because the department is not contained in the FASTDEPT group

- '020177,SALES   ' → MEDIUM, because the account number does not match group SLOWACCT, and therefore no sub-rules are checked

How WLM matches qualifier values

- When matching on qualifier values shorter than 8 characters, WLM treats long and short qualifier types differently:
  - **Short qualifier types:** Value padded with blanks to be 8 characters, blanks used for matching
  - **Long qualifier types with start position:** Value padded with blanks to be 8 characters, blanks used for matching
  - **Long qualifier types without start position:** Value matched according to the number of characters specified
- Example:

```
--------Qualifier--------                    -------Class--------
Type        Name        Start                Service     Report
                                  DEFAULTS:  MEDIUM      _____
AI          DIRS        8                    SLOW        _____
AI          DIRS*       8                    FAST        _____
AI          0201        ____                 FAST        _____
UI          HUGO        ____                 SLOW        _____
```

  - 1st rule matches accounting information with the 8 characters 'DIRS____' starting in the 8th position
  - 2nd rule matches accounting information with the 4 characters 'DIRS' starting in the 8th position
  - 3rd rule matches accounting information starting with the 4 characters '0201'
  - 4th rule matches user ID equal to the 8 characters 'HUGO____'

Coexistence and migration considerations for new classification qualifiers and groups

- Apply toleration APAR OA36842 to z/OS V1.10 through V1.13 to handle service definitions with functionality level 29 introduced by use of z/OS V2R1 functionality
  - Service definitions with functionality level 29 cannot be extracted, displayed, modified, installed or activated in a back-level WLM Administrative Application
  - But they can be activated in a mixed z/OS V1.10 through V2.1 Sysplex using
    - The WLM Administrative Application on the z/OS V2.1 system
    - Console command "v wlm,policy=<pol>" on the z/OS V2.1 system
    - WLM service IWMPACT on the z/OS V2.1 system
  - WLM then runs with this service definition on all systems
  - However, the new groups and new and modified qualifier types are **not** honored for workload classification on pre-z/OS V2.1 systems

Coexistence and migration considerations for new classification qualifiers and groups

- If you plan to use more than 500 lines of notepad information, re-allocate the WLM couple data set on the z/OS V2R1 system before installing the service definition

  - By using z/OS V2.1 to allocate the WLM couple data set, the space allocated is sufficient for the increased notepad size

  - Else you may receive error message "WLM couple data set is too small to hold the service definition. (IWMAM047)"

z/OS release Function	V2.1	V1.13 – V1.10
Groups of SPM rules & new classification qualifiers	**+**	*Toleration OA36842*

Agenda

- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- I/O Priority Groups
- Other z/OS V2.1 Enhancements
- Manage CICS Regions Using Goals Of: "BOTH"
- Response Time Distribution for Velocity Goals
- HiperDispatch Enhancements

I/O Priority Groups

- **Rationale**

    - I/O Priority is used to control DASD I/O queuing.

    - WLM dynamically adjusts the I/O priority based on goal attainment and whether the device can contribute to achieve the goal.

    - Every 10 minutes, WLM determines which service classes use which devices and builds so called device sets.

    - Typically, different workloads use distinct device sets and WLM changes I/O priorities between service classes using the same device set.

    - If a workload starts to use a device outside from its previously used device sets and experiences significant I/O delay, it may take up to 10 minutes until WLM refreshes the device sets and adapts the I/O priority of the corresponding service class.

- **Solution:**

    - Important service classes which are sensitive to I/O delay can now be assigned to priority group HIGH which ensures that they get always higher I/O priorities than the service classes assigned to group NORMAL.

Complete your sessions evaluation online at SHARE.org/BostonEval
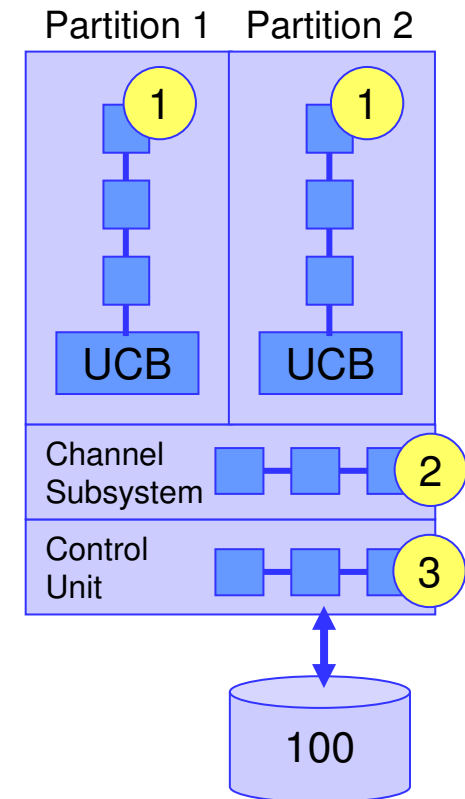
I/O Priority Groups Specification in WLM ISPF Application

z/OS release / Function	V2.1	V1.13	V1.12
I/O Priority Groups	**+**	**Toleration OA37824**	**Toleration OA37824**

Specify the I/O Priority Group in the WLM ISPF Administrative Application:

```
                       Create a Service Class


Command ===> _____
Service Class Name . . . . . _____       (Required)
Description  . . . . . . . . _____
Workload Name  . . . . . . . _____       (name or ?)
Base Resource Group  . . . . _____       (name or ?)
Cpu Critical . . . . . . . . NO_           (YES or NO)
I/O Priority Group . . . . . NORMAL        (NORMAL or HIGH)
```

I/O Priority Groups – Validation

Group HIGH is only honored by WLM if "I/O priority management" and "I/O priority groups" are enabled:

```
                  Service Coefficient/Service Definition Options

    I/O priority management  . . . . . . . NO_  (Yes or No)
    Enable I/O priority groups . . . . . . YES  (Yes or No)
    Dynamic alias tuning management  . . . NO_  (Yes or No)
```

The "Validate definition" option can be used to check whether service classes assigned to I/O priority group HIGH although I/O priority management is not enabled

```
                  Service Definition Validation Results

  IWMAM918W  Service class(es) assigned to I/O priority group HIGH but
  _____   I/O priority management or I/O priority groups are not
  _____   enabled. The I/O priority group will not be honored.
```

I/O Priority Groups – Specification in z/OSMF

z/OSMF Workload Management task provides new option, too.

Complete your sessions evaluation online at SHARE.org/BostonEval

in Boston

I/O Priority Groups – Callable Services

- The WLM services **IWMDEXTR** or **IWMDINST** allows extracting or installing a service definition in XML format.
- The layout of the XML service definition (DTD) is extended as follows. The entire DTD is described in Appendix C of the WLM Services Guide.

```
<!ELEMENT ServiceClass ( Name, Description?, CreationDate,
        CreationUser, ModificationDate, ModificationUser,
        CPUCritical?, IOPriorityGroup?, ResourceGroupName?,
        Goal ) >

<!ELEMENT ServiceClassOverride ( ServiceClassName,
        CPUCritical?, IOPriorityGroup?,ResourceGroupName?,
        Goal ) >

<!ELEMENT IOPriorityGroup ( #PCDATA ) >
```

```
<!ELEMENT ServiceOptions ( IOPriorityManagement,
DynamicAliasManagement?, IOPriorityGroupsEnabled? ) >

<!ELEMENT IOPriorityGroupsEnabled ( #PCDATA ) >
```

I/O Priority Groups – Callable services

- The RASD parameter list of **SYSEVENT REQASD** and **REQFASD** is extended to return information about the I/O priority group of the address space. Additional flags are added to field RASDFLAGS1.

- **IWMRQRY** is the interface reporting products should use to obtain address space related general execution delays. The answer area mapped by IWMWRQAA is enhanced according to REQFASD. An additional flag is added to field RQAEFLG1.

- **IWMPQRY** is the interface to return a representation of the active policy. The answer area mapped by IWMSVPOL is extended. An additional flag is added to SVPOLCFL of the service class definition section SVPOLC.

I/O Priority Groups – SMF record type 72.3

RMF's record types 72 subtype 3 and SMF 79 subtypes 1 and 2 are extended to indicate assignment to the I/O priority group.

SMF record 72 subtype 3 (Workload activity) – Workload manager control section				
Offsets	Name	Len	Format	Description
0 0	R723MSCF	1	Binary	Service/Report class flags. Bit 0-6: Meaning not changed **Bit 7: Indicator for I/O priority group HIGH**

SHARE in Boston

I/O Priority Groups – SMF record type 79

RMF's record types 72 subtype 3 and SMF 79 subtypes 1 and 2 are extended to indicate assignment to the I/O priority group.

SMF record 79 subtype 1 (Address space state data) – ASD data section				
Offsets	Name	Len	Format	Description
236 EC	R791FLG3	1	Binary	Additional flags. **Bit 0: Service class assigned by classification or RESET SRVCLASS belongs to I/O priority group HIGH in the active policy** **Bit 1: I/O priority group HIGH was assigned either to the address space or to transaction service classes served by the space** Bit 2-7: Reserved
SMF record 79 subtype 2 (address space resource data) – ARD data section				
224 E0	R792FLG3	1	binary	Additional flags. **Bit 0: Service class assigned by classification or RESET SRVCLASS belongs to I/O priority group HIGH in the active policy** **Bit 1: I/O priority group HIGH was assigned either to the address space or to transaction service classes served by the space** Bit 2-7: Reserved

I/O Priority Groups – RMF: Workload Activity Report

- Postprocessor Workload Activity (WLMGL) report is extended.
- If service class is assigned to I/O priority group HIGH, an indication is displayed in the SERVICE CLASS(ES) and SERVICE CLASS PERIODS sections.

```
----------------------------------------------------------------------------------------------- SERVICE CLASS(ES)

REPORT BY: POLICY=WLMPOL       WORKLOAD=ONLINE        SERVICE CLASS=ONLTOP      RESOURCE GROUP=*NONE
                                                      CRITICAL      =CPU+STORAGE                    I/O PRIORITY GROUP=HIGH
                                                      DESCRIPTION   =Batch Workload

-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---     SERVICE TIME   ---APPL %---   --PROMOTED--  ----STORAGE----
AVG      0.74   ACTUAL                 0  SSCHRT  0.0   IOC          0    CPU    6.429  CP      0.66   BLK    0.000  AVG     7663.01
MPL      0.74   EXECUTION              0  RESP    0.0   CPU     287332    SRB    0.000  AAPCP   0.00   ENQ    0.000  TOTAL   5698.61
ENDED       0   QUEUED                 0  CONN    0.0   MSO     537297    RCT    0.002  IIPCP   0.00   CRM    0.000  SHARED     0.00
```

I/O Priority Groups require some migration and coexistence considerations

- Toleration **APAR OA37824** required on z/OS V1R12 and z/OS V1R13 systems because dynamic I/O priority management is a sysplex-wide function
- Turn on I/O priorities only if all systems sharing disk systems run on z/OS V2R1 or on z/OS V1R12 / R13 with OA37824
- When the Enable I/O Priority Groups option is turned on in one sysplex, turn it also on in other sysplexes even if they do not exploit I/O priority group HIGH.
  - Ensures that all systems sharing a disk system work with an identical range of I/O priorities
- Assigning service classes to I/O priority group HIGH is only possible with the z/OS V2R1 WLM ISPF Application or z/OSMF V2R1
- If a service class is assigned to I/O priority group HIGH, the functionality level of the service definition is increased to **LEVEL029**
  - A service definition at functionality level 29 cannot be extracted, displayed, modified, installed, or activated by an WLM Application prior z/OS V2R1
- RMF support is only available with z/OS V2R1

Agenda

- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- I/O Priority Groups
- Other z/OS V2.1 Enhancements
  - Improved granularity for resource groups
  - 3000 Application Environments
- Manage CICS Regions Using Goals Of: "BOTH"
- Response Time Distribution for Velocity Goals
- HiperDispatch Enhancements

Improved granularity for resource group capping

- To enforce resource group capping dispatchable units are marked non-dispatchable or dispatchable (awake slice) for some time
- Smallest resource group limit and granularity that can be enforced depends on

1	9	17	25	33	41	49	57
2	10	18	26	34	42	50	58
3	11	19	27	35	43	51	59
4	12	20	28	36	44	52	60
5	13	21	29	37	45	53	61
6	14	22	30	38	46	54	62
7	15	23	31	39	47	55	63
8	16	24	32	40	48	56	64

- Processor speed/capacity
- Number logical processors in system or Sysplex, and number of dispatchable units of the work to be capped
- Service consumed at higher priority than capped work
- With z/OS V2.1 the number of time slices for resource group management was quadrupled
- From 1/64th to 1/256th of elapsed time
- Allows for more fine grain control of resource groups

More than 999 Application Environments

- A –static- application environment is a named entity in the WLM service definition that allows WLM to start server address spaces for scalable client/server type applications.

  - One of the main exploiters of this function are DB2 Stored Procedures

- Large DB2 installations may have a requirement to define more than 999 static Application Environments

  - Typically, these are SAP installations where the WLM service definition is shared across many systems and Sysplexes

- With z/OS V2.1 WLM increases the limit from 999 to 3000.

More than 999 Application Environments: Coexistence considerations

- As soon as more than 999 AEs are defined, the functionality level of the service definition is raised to **LEVEL029**
  - Can use z/OS V2.1 WLM ISPF Application or z/OSMF V2.1
  - Any service definition at functionality level 29 cannot be extracted, displayed, modified, installed, or activated by an WLM ISPF Application prior z/OS V2R1
  - If a service definition at LEVEL029 is installed to the WLM Couple Data Set by z/OS V2R1, systems with z/OS V1R12 and V1R13 of the same Sysplex can activate the policy

- APAR OA36842 for toleration of level 29 policies required on z/OS V1R12 and V1R13

More than 999 Application Environments: Migration considerations

- Customers with a need for more than 999 AEs must allocate a Couple Data Set for WLM which can hold the required number of AE objects
  - This is achieved by performing the **Allocate couple data set using CDS values** task in the WLM ISPF application
    - If a service definition with more AEs than allowed for the current WLM couple data set would be installed, the WLM ISPF application displays message
    IWMAM047 WLM couple data set is too small to hold the service definition

  - Alternatively, it is possible to allocate a WLM couple data set by running a job as provided in SYS1.SAMPLIB(IWMFTCDS)

Agenda

- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- I/O Priority Groups
- Other z/OS V2.1 Enhancements
- Manage CICS Regions Using Goals Of: "BOTH"
- Response Time Distribution for Velocity Goals
- HiperDispatch Enhancements

Manage Regions Using Goals Of: "BOTH":
New management option for CICS environments

- **Problem:**
  In environments with pre-dominant CICS workloads it is possible to observe contention problems as described in the following sample test scenario

- **Example:**
  Workload: Websphere → CICS → DB2
  - Websphere receives work, sends it to CICS TORs which send it to AORs which execute DB2 calls
  - Classification: Websphere Imp=2 and all CICS Imp=3, managed towards response time goals
- *Symptoms:*
  - Low system throughput
  - Relatively high response times
  - System utilization can hardly be increased beyond (in this scenario) 80%

Manage Regions Using Goals Of: "BOTH"

CPU Service Consumption by Importance Level
Test Run: P9185

CICS Workload: All TORs and all AORs

WAS workload: Each WAS Transaction is generating a CICS transactions

Legend: CPU System, CPU Imp=1, CPU Imp=2, CPU Imp=3, CPU Imp=4, CPU Imp=5, CPU Disc, CPU Free
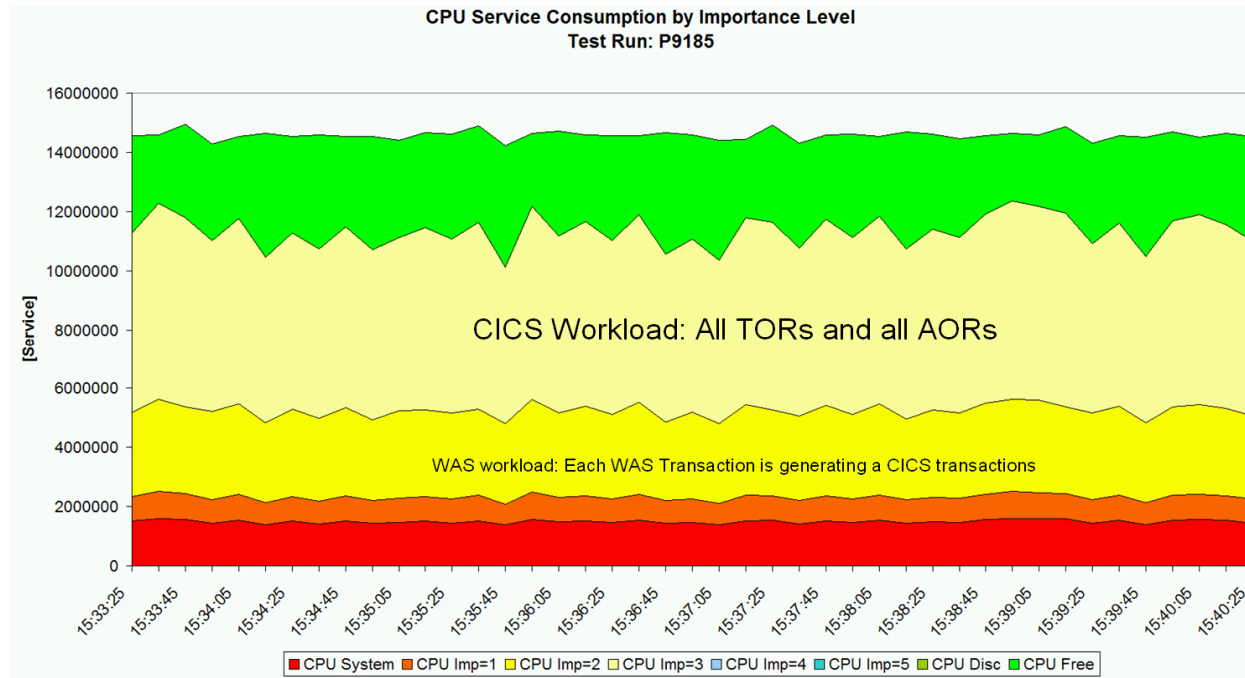
- Problem Analysis
  - TORs and AORs run at the same dispatch priority
  - AORs heavily consume CPU.
    TORs compete against AORs and need to wait too long to receive work and return results to the caller fast enough
  - Hiperdispatch can amplify the situation because it runs the work at higher utilization

Manage Regions Using Goals Of: "BOTH"
Sample test results
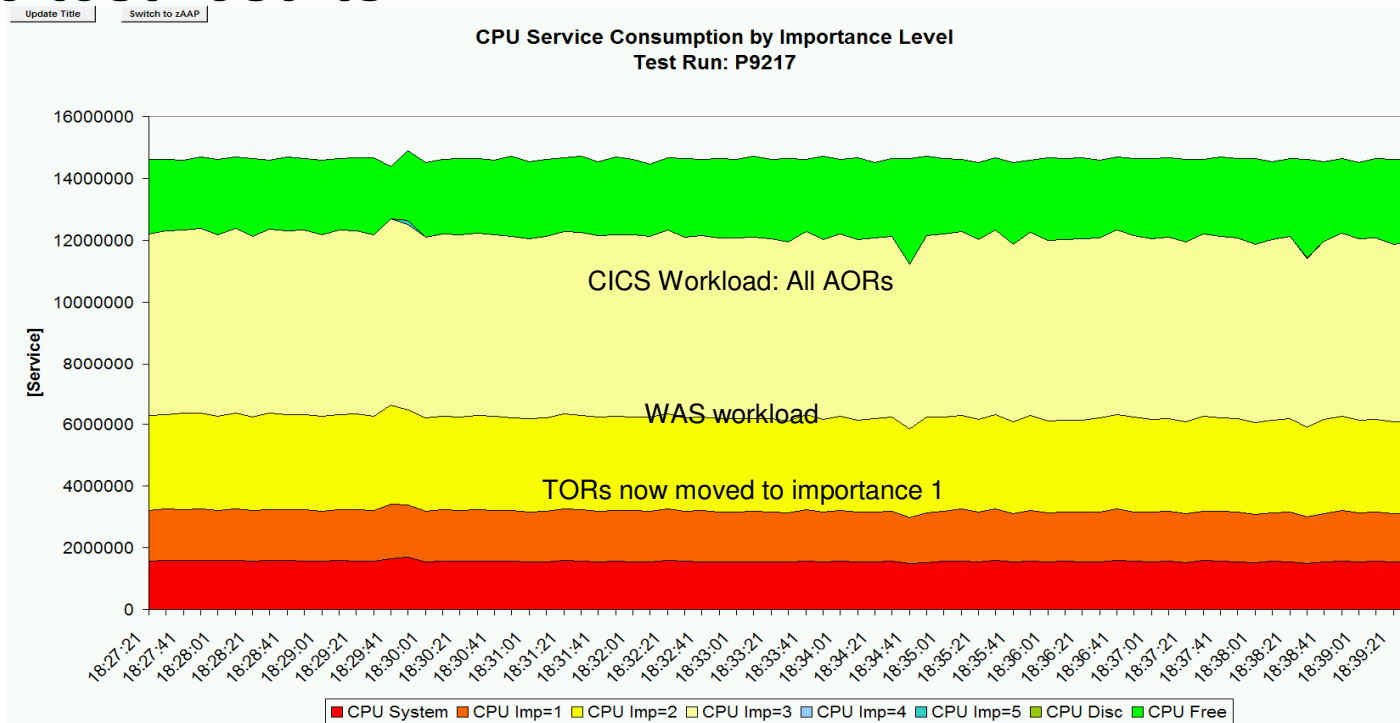
Test example:

	Completed Transactions/sec	Avg. RespTime/sec
w/o „BOTH"	9765	0.197
„BOTH"	12463 **+27%**	0.026 **-86%**

Throughput enhancement and significant response time reduction.

Manage Regions Using Goals Of: "BOTH"

- **Two possible alternatives**:
  (Move TORs to a service class with higher importance than AORs)

  - Option 1: Exempt **all** regions from being managed by response time goals and classify TORs to a service class with higher importance than AORs.

    Disadvantage: No response time data present

  - Option 2: Exempt only AORs and move them to a service class with lower importance than the CICS service classes with response time goals.

    Disadvantage: Response time data cover only a small portion of the execution path because AORs consume much more than TORs.

Manage Regions Using Goals Of: "BOTH"

```
 Subsystem-Type  Xref  Notes  Options  Help
----------------------------------------------------------------
             Modify Rules for the Subsystem Type      Row 1 to 3 of 3
Command ===> _____       Scroll ===> PAGE

Subsystem Type . : JES         Fold qualifier names?   Y  (Y or N)
Description   . . . Batch Work_____

Action codes:     A=After       C=Copy        M=Move        I=Insert rule
                  B=Before      D=Delete row  R=Repeat      IS=Insert Sub-rule
                                                            <=== More
          --------Qualifier--------          Storage    Manage Region
Action    Type      Name      Start          Critical   Using Goals Of

_____  1  TN       CICSTOR* ___            NO         BOTH
_____  1  TN       CICSAOR* ___            NO         TRANSACTION   ■
_____  1  TN       CICS*    ___            NO         TRANSACTION
***************************** BOTTOM OF DATA ******************************
```

- Define STC service class for TORs which has a higher importance than the CICS service class with response time goals for the CICS work and AORs
- TOR: "Manage Regions by Goals Of: BOTH" option in WLM service definition
- AOR: "Manage Regions by Goals Of: TRANSACTION" (the default)

- Result:
  - WLM will manage the TORs towards the goals of the STC service class
  - <u>And</u> WLM will ensure bookkeeping of transaction completions to the correct CICS response time service class
  - The CICS transactions are managed towards CICS response time goals and the AORs are also managed towards these goals like today

Agenda

- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- 3000 Application Environments
- I/O Priority Groups
- Manage CICS Regions Using Goals Of: "BOTH"
- Response Time Distribution for Velocity Goals
- HiperDispatch Enhancements

Response Time Distribution for Velocity Goals

- Before this support, WLM reporting did not provide a response time distribution (ended transactions) for workloads with velocity goals

- Sometimes it is desirable to have a response time distribution for <u>all</u> transactional workloads, even if they have a velocity goal
  - More data to analyze workload behavior and to detect problems
  - Better support for migration of goal definitions to response time goals

- With z/OS V1.13
  - the IWMRCOLL answer area IWMWRCAA provides also a response time distribution for service class periods with an execution velocity goal
  - the RMF Postprocessor Workload Activity report (WLMGL) displays the new response time distributions

SHARE
in Boston

Response Time Distribution for Velocity Goals

Th

Response Time Distribution

	<332	<=398	<=465	<=531	<=598	<=664	<=731	<=797	<=864	<=930	<=996	<=1329	<=2658	>2658
#Transactions	4	0	0	1	0	6	4	2	0	0	1	1	0	0

RT (sec)
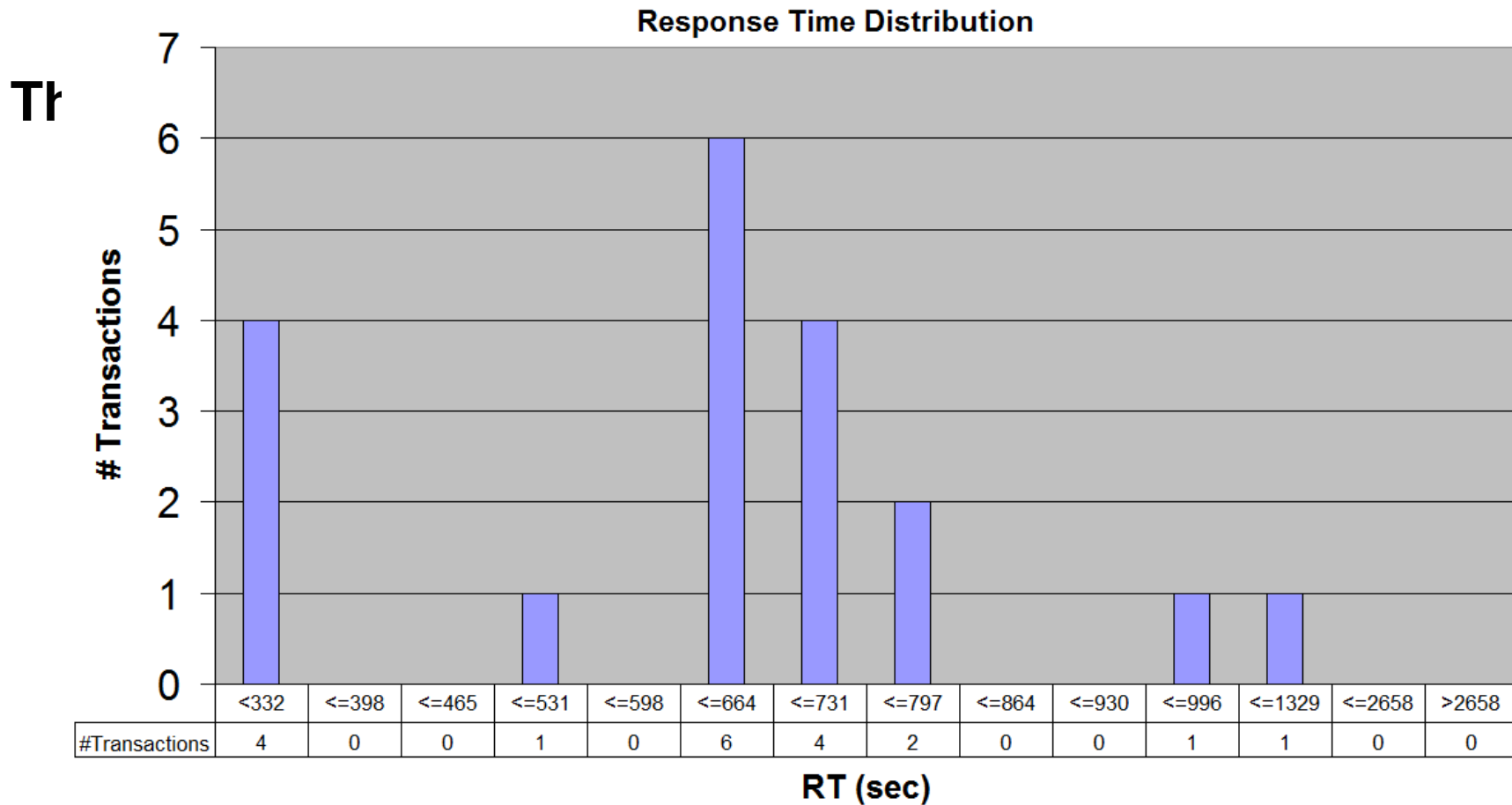
- Response Time goals use the goal as „reference" (bucket 6)
- Velocity goals do not have a „reference" response time
- The mid-point (MP) is calculated based on workload

Response Time Distribution for Velocity Goals

RMF Postprocessor Workload Activity Report:

```
REPORT BY: POLICY=DAYPOL      WORKLOAD=BATCH        SERVICE CLASS=SBATPMD     RESOURCE GROUP=*NONE      PERIOD=1 IMPORTANCE=4
                                                    CRITICAL     =NONE


-TRANSACTIONS-   TRANS-TIME HHH.MM.SS.TTT   --DASD I/O--   ---SERVICE---   SERVICE TIME   ---APPL %---   --PROMOTED--   ----STORAGE----
AVG      1.70    ACTUAL          4.06.826   SSCHRT  1467   IOC   133593   CPU   64.577   CP      8.94   BLK   0.000   AVG     1217.86
MPL      1.70    EXECUTION         28.632   RESP     0.1   CPU     1770K  SRB    8.644   AAPCP   0.00   ENQ   1.480   TOTAL   2074.61
ENDED      44    QUEUED             1.639   CONN     0.0   MSO        0   RCT    0.007   IIPCP   0.00   CRM   0.000   SHARED    11.66
END/S    0.05    R/S AFFIN              0   DISC     0.1   SRB   236535   IIT    7.204                  LCK   1.617
#SWAPS      2    INELIGIBLE      3.36.554   Q+PEND   0.1   TOT     2140K  HST    0.000   AAP     N/A    SUP   0.000   -PAGE-IN RATES-
EXCTD       0    CONVERSION            47   IOSQ     0.0   /SEC    2378   AAP     N/A    IIP     0.00                  SINGLE      0.0
AVG ENC  0.00    STD DEV         5.35.053                                IIP    0.000                                BLOCK       0.0
REM ENC  0.00                                              ABSRPTN  1396                                             SHARED      0.0
MS ENC   0.00                                              TRX SERV 1396                                             HSP         0.0


GOAL: EXECUTION VELOCITY 20.0%     VELOCITY MIGRATION:    I/O MGMT  81.2%     INIT MGMT 78.2%

          RESPONSE TIME EX   PERF   AVG    --EXEC USING%--   -------------- EXEC DELAYS % -----------   -USING%-  --- DELAY % ---    %
SYSTEM                 VEL% INDX ADRSP   CPU AAP IIP I/O   TOT CPU I/O                                  CRY CNT   UNK IDL CRY CNT  QUI

*ALL       --N/A--     81.2  0.2   1.8   2.8 N/A 0.0  50   12 9.2 3.0                                  0.0 0.0    33 2.7 0.0 0.0  0.0
SYS1                   82.9  0.2   0.7   3.5 N/A 0.0  75   16  15 1.6                                  0.0 0.0   2.2 3.4 0.0 0.0  0.0
SYS2                   71.2  0.3   1.1   1.7 N/A 0.0  14  6.5 1.5 4.9                                  0.0 0.0    76 1.8 0.0 0.0  0.0


                         ----------RESPONSE TIME DISTRIBUTIONS----------

 SYSTEM: SYS1 ----INTERVAL: 00.15.00.000 ---MRT CHANGES:  0 ---    SYSTEM: SYS2 ----INTERVAL: 00.15.00.000 ---MRT CHANGES:  0 ---

    ---TIME----  -NUMBER OF TRANSACTIONS-  ------PERCENT------       ---TIME----  -NUMBER OF TRANSACTIONS-  ------PERCENT------
    HH.MM.SS.TTT CUM TOTAL     IN BUCKET   CUM TOTAL IN BUCKET       HH.MM.SS.TTT CUM TOTAL     IN BUCKET   CUM TOTAL IN BUCKET
 <  00.00.16.994        21           21        84.0       84.0    <  00.05.32.317         4            4        21.1       21.1
 <= 00.00.20.393        22            1        88.0        4.0    <= 00.06.38.781         4            0        21.1        0.0
 <= 00.00.23.792        24            2        96.0        8.0    <= 00.07.45.244         4            0        21.1        0.0
 <= 00.00.27.191        24            0        96.0        0.0    <= 00.08.51.708         5            1        26.3        5.3
 <= 00.00.30.590        24            0        96.0        0.0    <= 00.09.58.171         5            0        26.3        0.0
 <= 00.00.33.989        25            1         100        4.0    <= 00.11.04.635        11            6        57.9       31.6
 <= 00.00.37.387        25            0         100        0.0    <= 00.12.11.098        15            4        78.9       21.1
 <= 00.00.40.786        25            0         100        0.0    <= 00.13.17.562        17            2        89.5       10.5
 <= 00.00.44.185        25            0         100        0.0    <= 00.14.24.025        17            0        89.5        0.0
 <= 00.00.47.584        25            0         100        0.0    <= 00.15.30.489        17            0        89.5        0.0
 <= 00.00.50.983        25            0         100        0.0    <= 00.16.36.952        18            1        94.7        5.3
 <= 00.01.07.978        25            0         100        0.0    <= 00.22.09.270        19            1         100        5.3
 <= 00.02.15.956        25            0         100        0.0    <= 00.44.18.540        19            0         100        0.0
 >  00.02.15.956        25            0         100        0.0    >  00.44.18.540        19            0         100        0.0
```

IWMRCOLL enhancements for Service and Report Class Periods

Section	Field	Response time goals	Execution velocity goals
RCAEIHDR (RCAE period header)	**RCAEIMID** (mid-point in milliseconds)	Same as goal value (milliseconds)	0 after policy activation/refresh/IPL New value computed when WLM detects that current workload distribution deviates **too much** from RCAEIMID for a **too long** time
	RCAEIRCT (running count)	N/A (value always 0)	Total number of RCAEIMID changes since last policy activation
	RCAEITST (timestamp of last change)	Policy activation time	Time of last RCAEIMID change or time of last policy activation
RCAEDIST	**RCAEDENT**	No change Distribution centered around goal value	Centered around RCAEIMID Reset after each RCAEIMID change

- RCAEIRCT
  - Is reset to 0 after each policy activation/refresh/IPL
  - Is incremented each time the report class period becomes heterogeneous (when RCAEPLSC and RCAEPMCI are updated)
  - Is incremented each time a transaction is reported with a new mid-point/timestamp
- The 14 buckets of the report class period's response time distribution are reset to 0 when RCAEIRCT is updated
- RCAEIMID is copied from the current service class period's RCAEIMID each time the report class period's RCAEIRCT is incremented
- RCAEITST is copied from the current service class period's RCAEITST each time the report class period's RCAEIRCT is incremented

SHARE
in Boston

Agenda

- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- I/O Priority Groups
- Other z/OS V2.1 Enhancements
- Manage CICS Regions Using Goals Of: "BOTH"
- Response Time Distribution for Velocity Goals
- HiperDispatch Enhancements

OA42185: Uneven HiperDispatch Balancing on VERY large systems

- Symptom: On very large systems HiperDispatch may utilize affinity nodes unevenly:
  - Some nodes may be "overloaded" showing an MVS busy near 100%, others have plenty of free capacity.
  - Can typically be seen only with very stable workloads utilizing ≥ 50 CPs.
- WLM HiperDispatch balancing algorithms are responsible for assigning work to affinity nodes every 2 sec.
- z/OS dispatcher responsibility is to schedule the work units on the logical processors that make up the affinity node. Overloaded processors can receive help from less loaded processors.

MVS Busy by LCP# and AN

- The WLM algorithms were changed to more aggressively balance the work unit as the number of processors grow - resulting in a more even distribution of work across nodes.
  - Note: It is perfectly fine to see uneven processor utilization within nodes, for I/O enabled processors and also some uneven utilization across nodes (without overloading nodes, though)
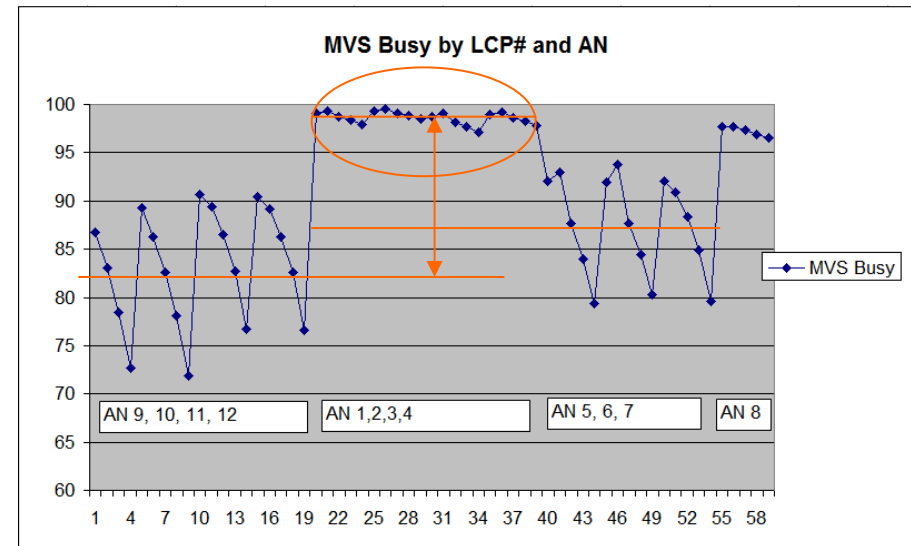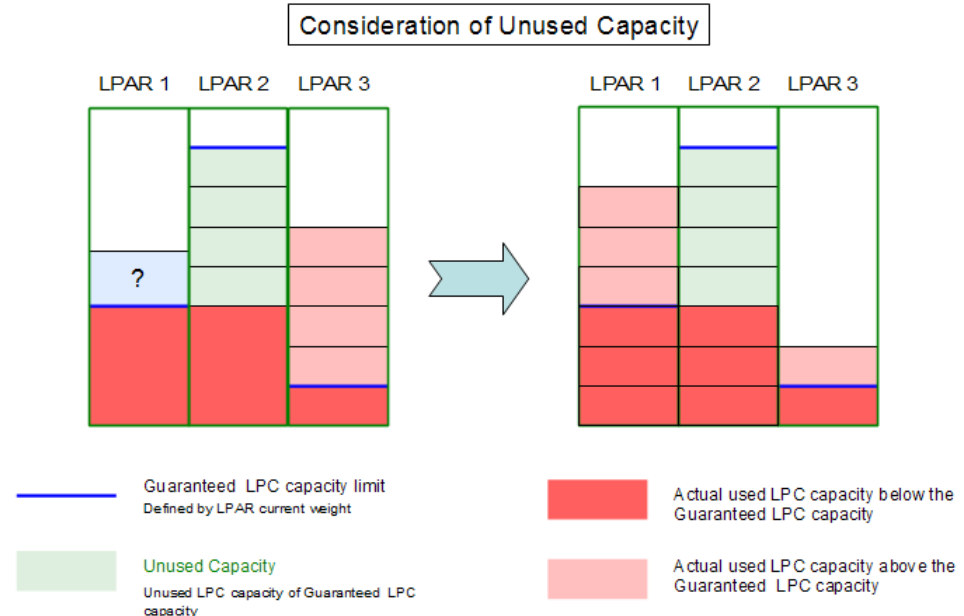
HiperDispatch Enhancements for "Unused Capacity"
OA37736 (z/OS V1R12, z/OS V1R13)

- Problem addressed:
  - If a large LPAR consumes below its weight-entitlement it is possible that a low-weight LPAR unparks many Vertical Low (VL) processors
  - Therefore a small LPAR could "dominate" larger LPARs because those could not unpark their VLs (additional VL would appear to be inefficient)

- Solution:
  - HiperDipatch considers now also the "unused" capacity share for a partition to unpark VLs
    - This share is calculated by dividing the unused capacity (guaranteed but not used) of all partitions in the CEC by the share of the partitions which can use more capacity

Consideration of Unused Capacity

LPAR 1 LPAR 2 LPAR 3   LPAR 1 LPAR 2 LPAR 3

—— Guaranteed LPC capacity limit
Defined by LPAR current weight

Unused Capacity
Unused LPC capacity of Guaranteed LPC capacity

Actual used LPC capacity below the Guaranteed LPC capacity

Actual used LPC capacity above the Guaranteed LPC capacity

SHARE in Boston

z/OS Workload Management
- More Information -

Workload Manager
Welcome to WLM/SRM
Overview   What's New   FAQs   Further Information

- z/OS WLM Homepage:

  http://www.ibm.com/systems/z/os/zos/features/wlm/

- z/OS MVS documentation
  - z/OS MVS Planning: Workload Management:
    http://publibz.boulder.ibm.com/epubs/pdf/iea2w1c0.pdf
  - z/OS MVS Programming: Workload Management Services:
    http://publibz.boulder.ibm.com/epubs/pdf/iea2w2c0.pdf
- *IBM Redbooks publication*s:
  - System Programmer's Guide to: Workload Manager:
    http://publib-b.boulder.ibm.com/abstracts/sg246472.html?Open
  - ABCs of z/OS System Programming Volume 12
    http://publib-b.boulder.ibm.com/abstracts/sg247621.html?Open

SHARE in Boston