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

Horst Sinram - STSM, z/OS Workload and Capacity Management
IBM Germany Research & Development

August 2014
Session 15714

#SHAREorg

SHARE in Pittsburgh 2014

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”
- Service Stream Enhancements
- WLM Managed DB2 Bufferpools

zAAP on zIPP (ZAAPZIIP) Changes

- In July 2013 IBM has modified the possible ratio of zIIP/zAAPs to CPs to be 2:1 for a zEC12 and/or zBC12.

- Customers may purchase up to two zIIP and/or up to two zAAP processors for every general purpose processor they purchase on the server. For the z196 and z114 and earlier servers, the 1:1 ratio will still be enforced; customers may purchase one zIIP and/or one zAAP for every general purpose processor they purchase on the server.
  - Refer to FAQs at http://www.ibm.com/systems/z/hardware/zenterprise/zec12.html

- SRM changed to relief requirements for the IEAOPTxx ZAAPZIIP ("zAAP on zIIP") option:
  - No longer limits the zAAP on zIIP function based on the number of zAAPs and/or the number of zIIPs installed on the machine.
    - While the *Global Performance Data Control* LPAR security setting is no longer required for this function GPDC continues to be required for other functions, namely HiperDispatch.
  - The zAAP on zIIP function continues to be limited to LPARs that have no zAAP.

Function / ***z/OS release***	*V2.1*	*V1.13*	*V1.12*
Relaxed ZAAPZIIP requirements	**+ *(OA43065 included in GA level)***	OA43102	OA43102

IBM zEnterprise EC12 GA2 Support Overview

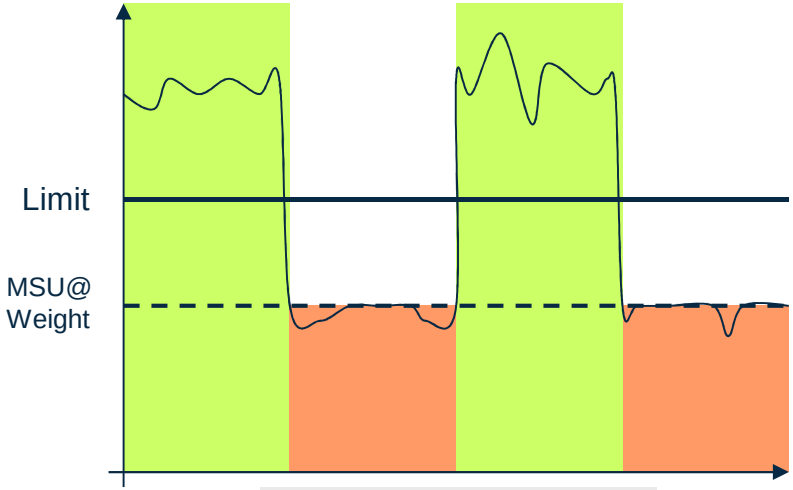
- zEnterprise BC12 and EC12 (zEC12) GA2 (firmware driver 15) offer new functions for hard and soft capping:
  - Smoother capping with WLM managed softcapping
  - When IRD weight management is active the group capacity of an LPAR may be derived by the initial weight
  - New “Absolute Capping Limit” LPAR 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 for defined capacity prior to zEC12 GA2

Pattern capping

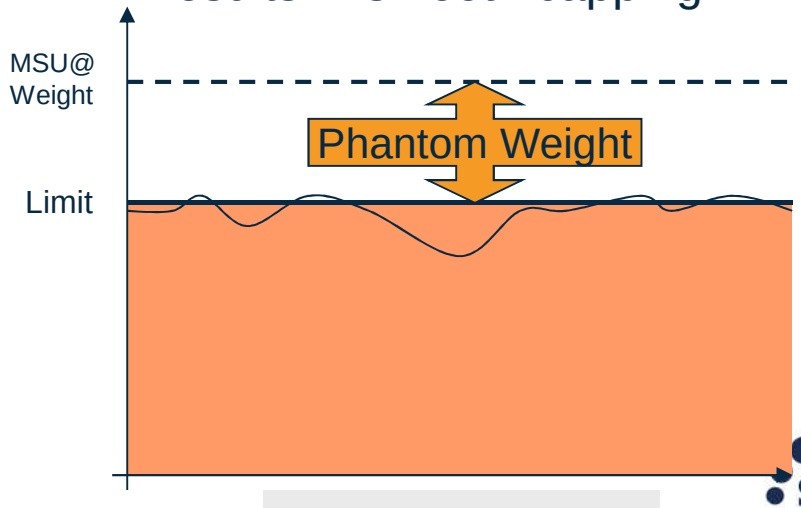
- Must be used when *MSU@LPARweight* < definedLimit
- Periods with LPAR capped at weight and running uncapped
- Can result in “pulsing” potentially impacting online workloads



MSU@Weight < Limit

Phantom weight capping

- Is used when *MSU@LPARweight* ≥ 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

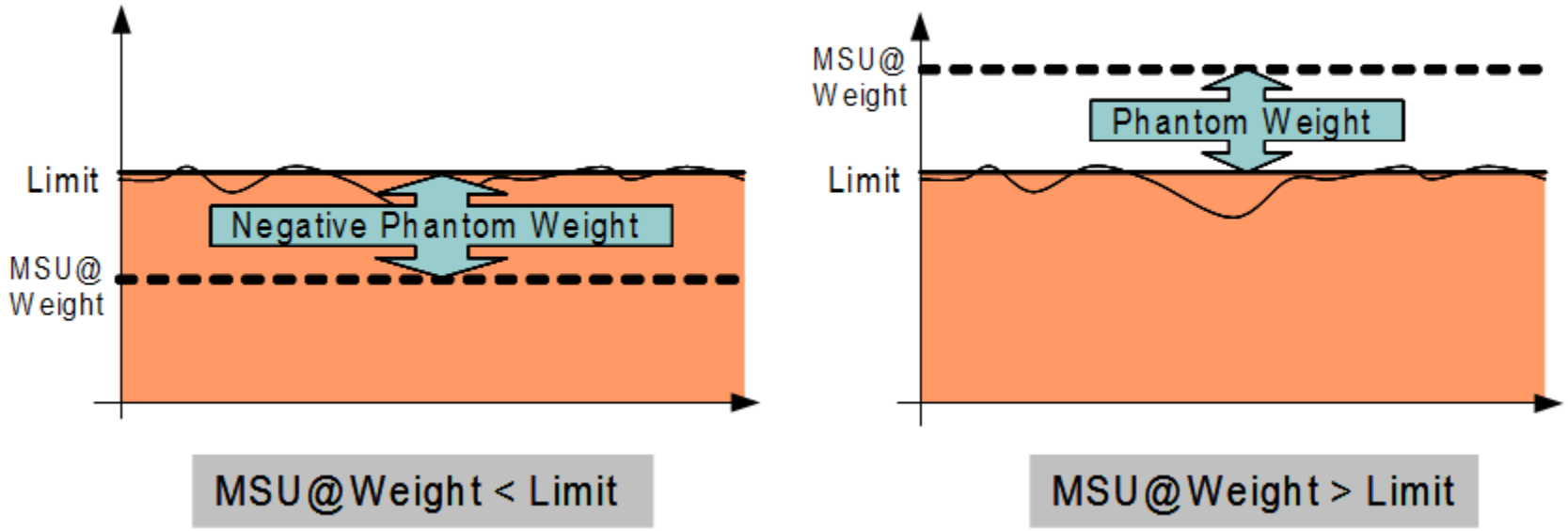


MSU@Weight > Limit

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

Capping Algorithm with Negative Phantom Weight



While a positive phantom weight changes the priority of a partition, the negative phantom weight caps the partition at a higher defined capacity without changing the priority of the partition.

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 run
    - z/OS V2.1, or
    - z/OS V1.12, V1.13 with OA41125 applied.
  - Results in more predictive and better controllable group capacity entitlement

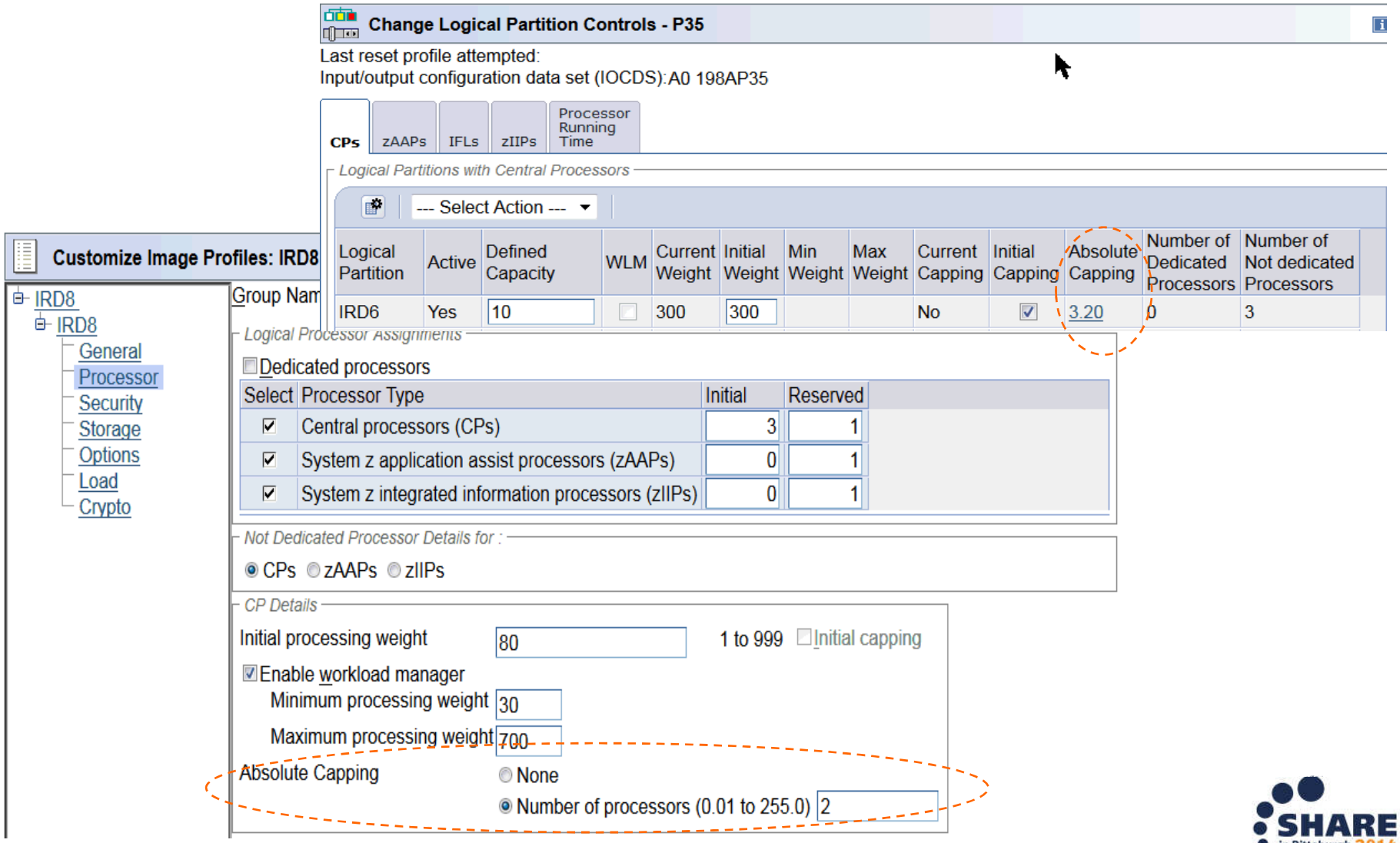
zEC12 GA2 Absolute Capping Limit

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

- Unlike initial capping it may be used *concurrently* with defined capacity and/or group capacity management

  - The minimum of all specified limits will be used
  - 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



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”
- Service Stream Enhancements
- WLM Managed DB2 Bufferpools

New Classification Qualifiers and Groups: Overview

- With z/OS V2R1, WLM/SRM introduces
  - New types of classification groups, and
  - Some new and modified types of work qualifiers for use in classification rules in the WLM service definition
- Can be used to improve the structure of your WLM service definition when masking or wild-carding are not sufficient to simplify classification rules.
- 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 to 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)

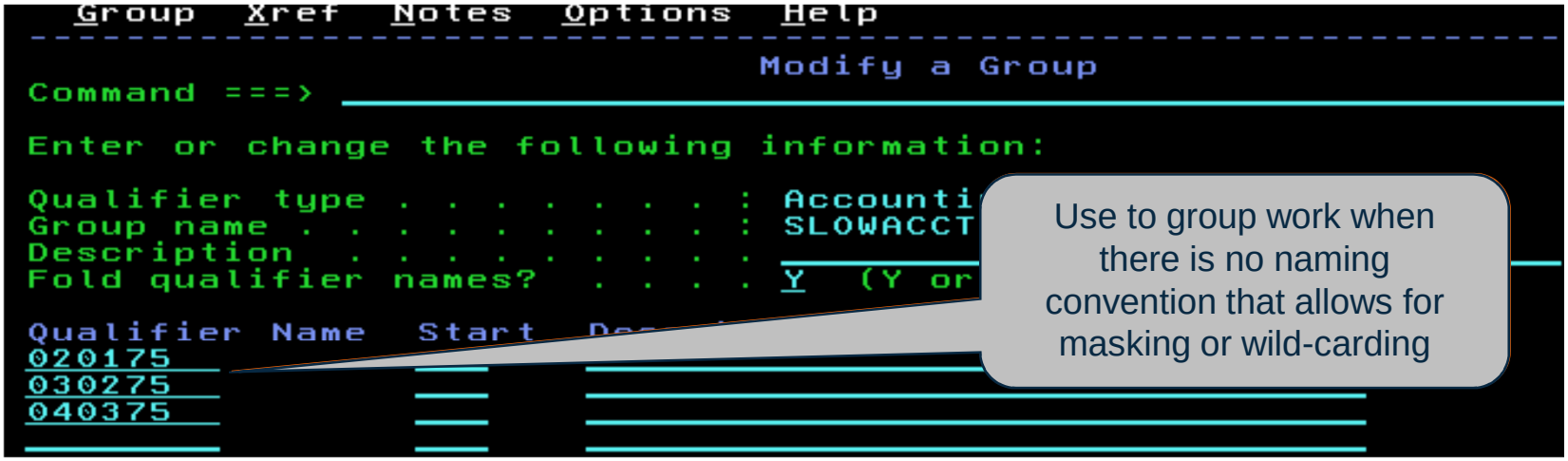
- The maximum number of “Notepad” lines the has been increased from 500 to 1000 lines

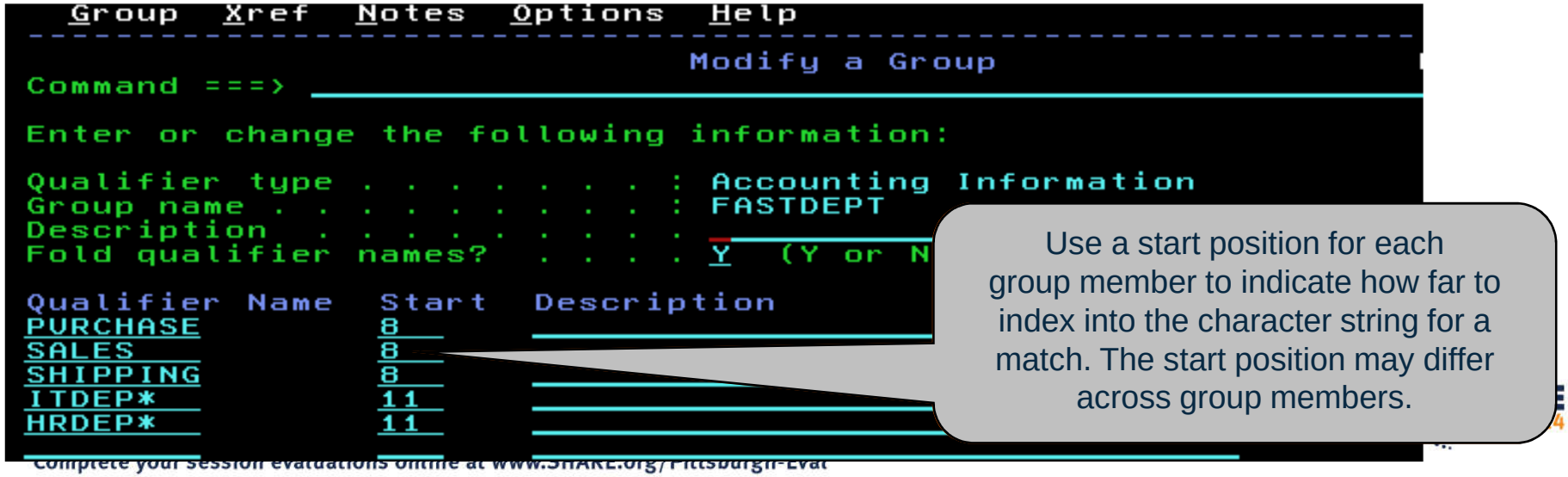
- Note: New and modified work qualifier types are only supported by the new 64-bit classify IWM4CLSY (planned to be used by DB2 V11).

WLM ISPF application enhancements

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

WLM ISPF application samples



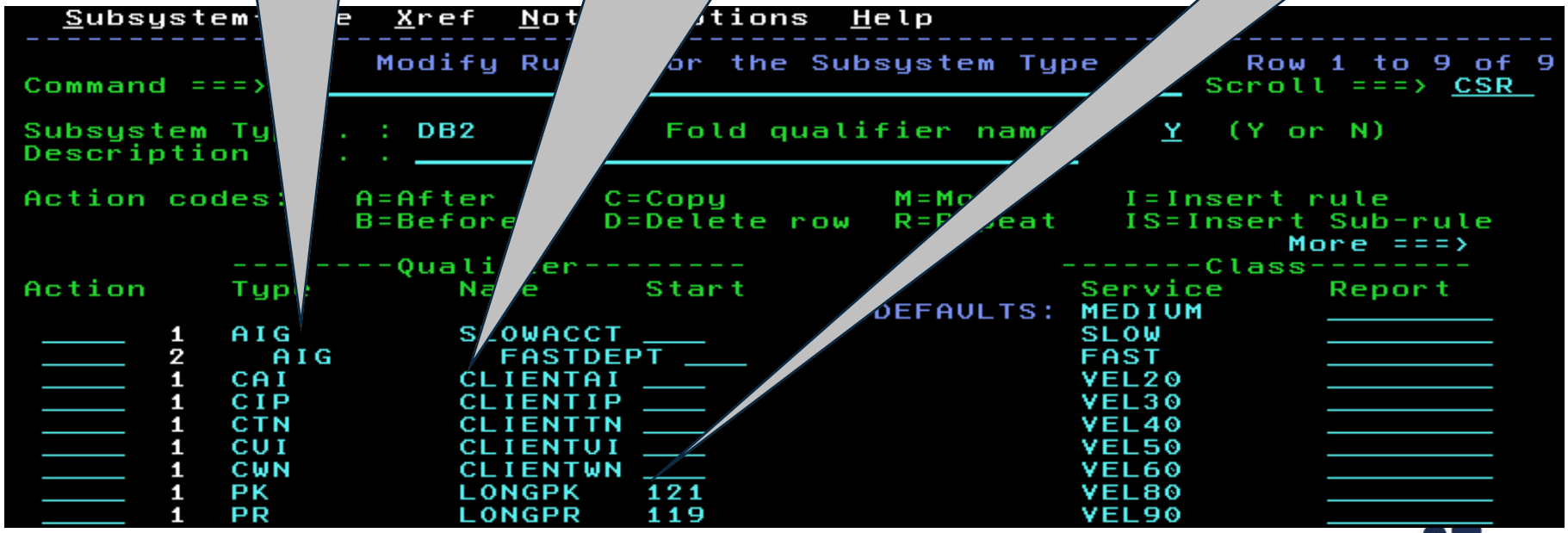


Use of New Groups and Qualifiers in the WLM Administrative Application

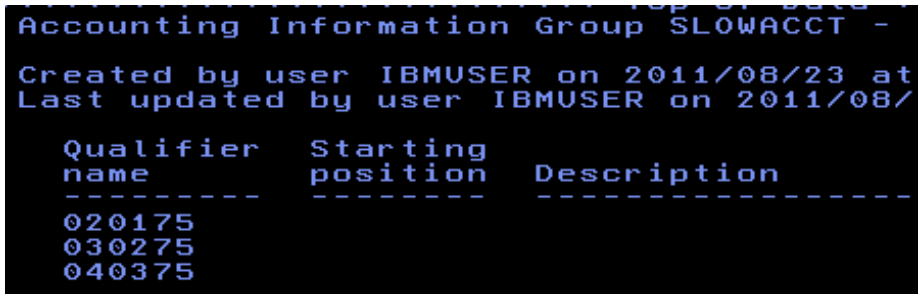
Groups of long work qualifier types can be nested

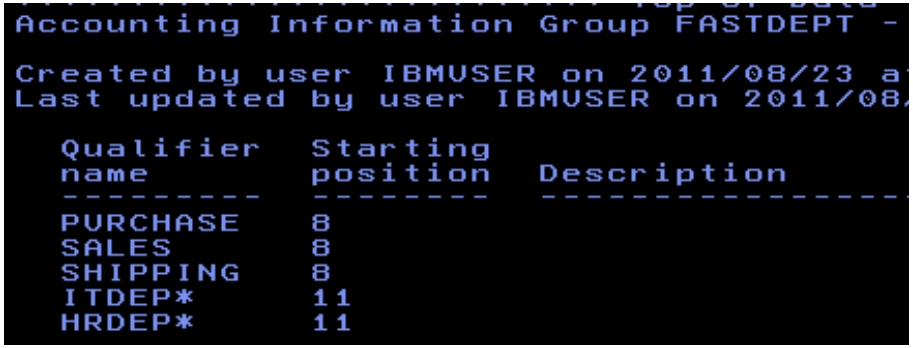
New work qualifier types:
- Client Accounting Information
- Client IP Address
- Client Transaction Name
- Client User ID
- Client Workstation Name

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



Classification via new groups: Examples





- '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)”

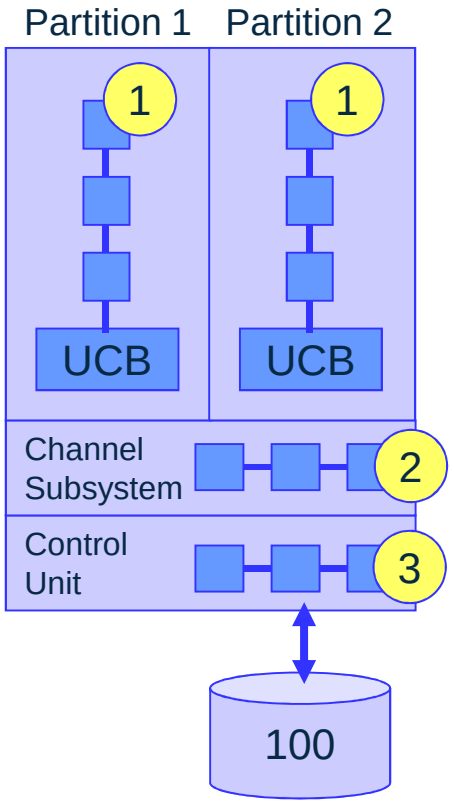
Function / ***z/OS release***	*V2.1*	*V1.13 – V1.10*
Groups of SPM rules & new classification qualifiers	**+**	***Toleration OA36842***

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”
- Service Stream Enhancements
- WLM Managed DB2 Bufferpools

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.

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

I/O Priority Group is specified in the service class definition:

```
                    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

But I/O Priority Group HIGH is only honored by WLM if both "I/O priority management" and "I/O priority groups" are enabled for the service definition:

```
    Service Coefficient/Service Definition Options

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

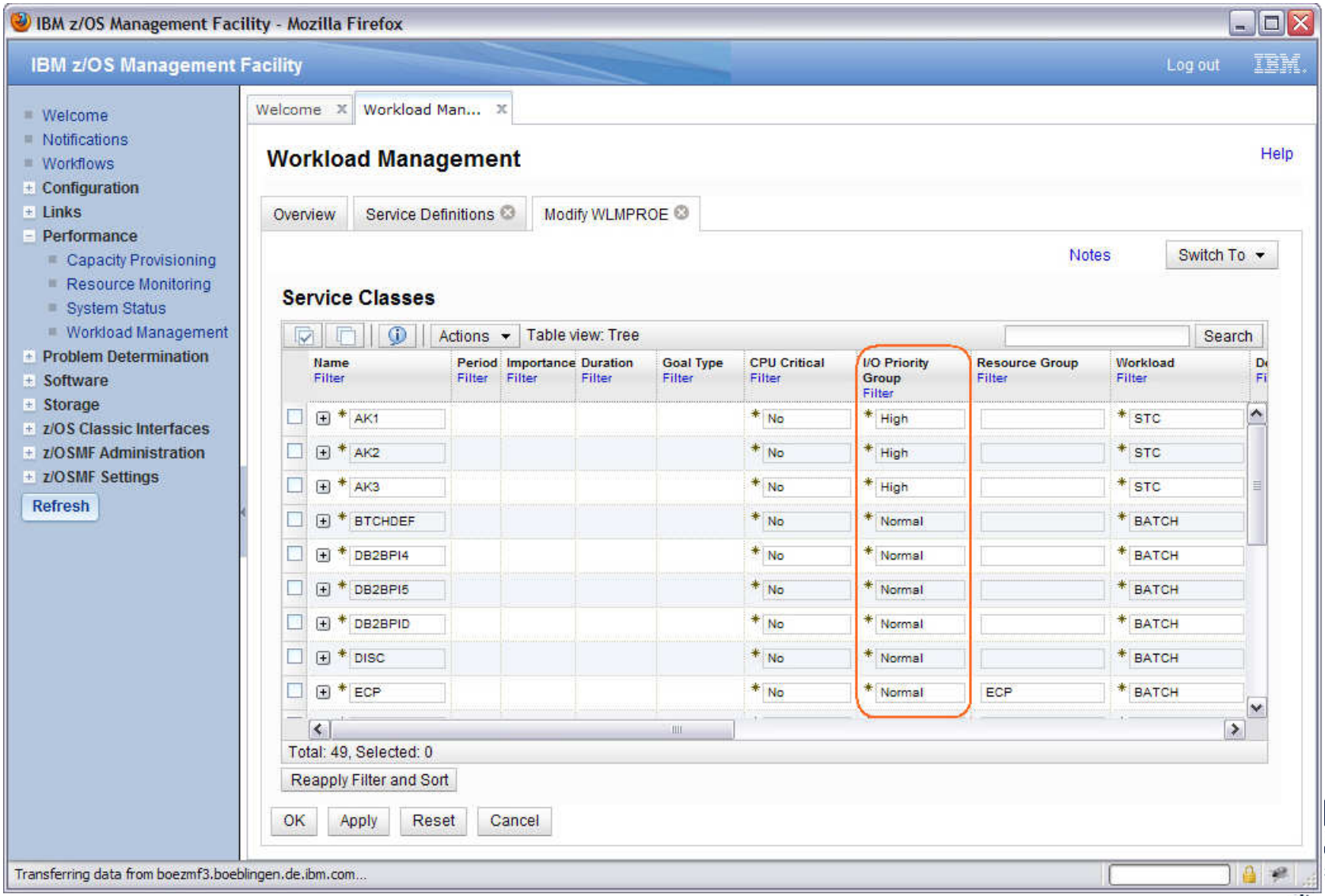
The "Validate definition" option can be used to check whether service classes assigned to I/O priority group HIGH although I/O priority management is not enabled

```
          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.



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**

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 shows I/O priority group
- 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
                                                  DESCRIPTION  =Batch Workload
                                                                                I/O PRIORITY GROUP=HIGH

-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
```

Use of I/O Priority Ranges

	I/O Priority Management=YES	
Priority	**I/O PriorityGroups NOT enabled**	**I/O PriorityGroup enabled**
FF	**SYSTEM**	**SYSTEM**
FE	**SYSSTC**	**SYSSTC**
FD	**Dynamically managed**	**Priority Group = HIGH**
FC		
FB		
FA		
F9		
F8		**Priority Group = NORMAL**
F7		
F6		
F5		
F4		
F3		
F2	**Discretionary**	**Discretionary**

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

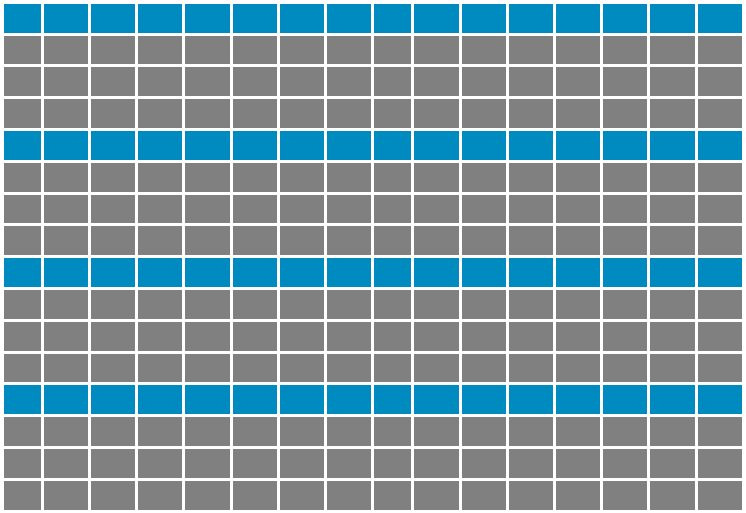
- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- I/O Priority Groups
- Other z/OS V2.1 Enhancements
  - Improved granularity for resource groups
  - 3000 Application Environments
- Manage CICS Regions Using Goals Of: “BOTH”
- Service Stream Enhancements
- WLM Managed DB2 Bufferpools

How Resource Groups (RG) Limit the CPU Consumption

- When WLM recognizes that the workload associated to a resource group exceeds its maximum limit, it will reduce processor access access by capping the amount of service that can be consumed by the RG.
- To do the capping, WLM calculates a pattern of intervals where the workload is set to *non-dispatchable* in some of the intervals.
- All service classes which belong to a RG are *awake* at the same time and they are also non-dispatchable at the same time.
- The RG cap pattern defines the time slices during which the workload is

Improved Granularity for Resource Group Capping in z/OS V2.1

- Smallest resource group limit and granularity that can be enforced depends on
  - Processor speed/capacity
  - Number of logical processors in system or Sysplex
  - 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 grained 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

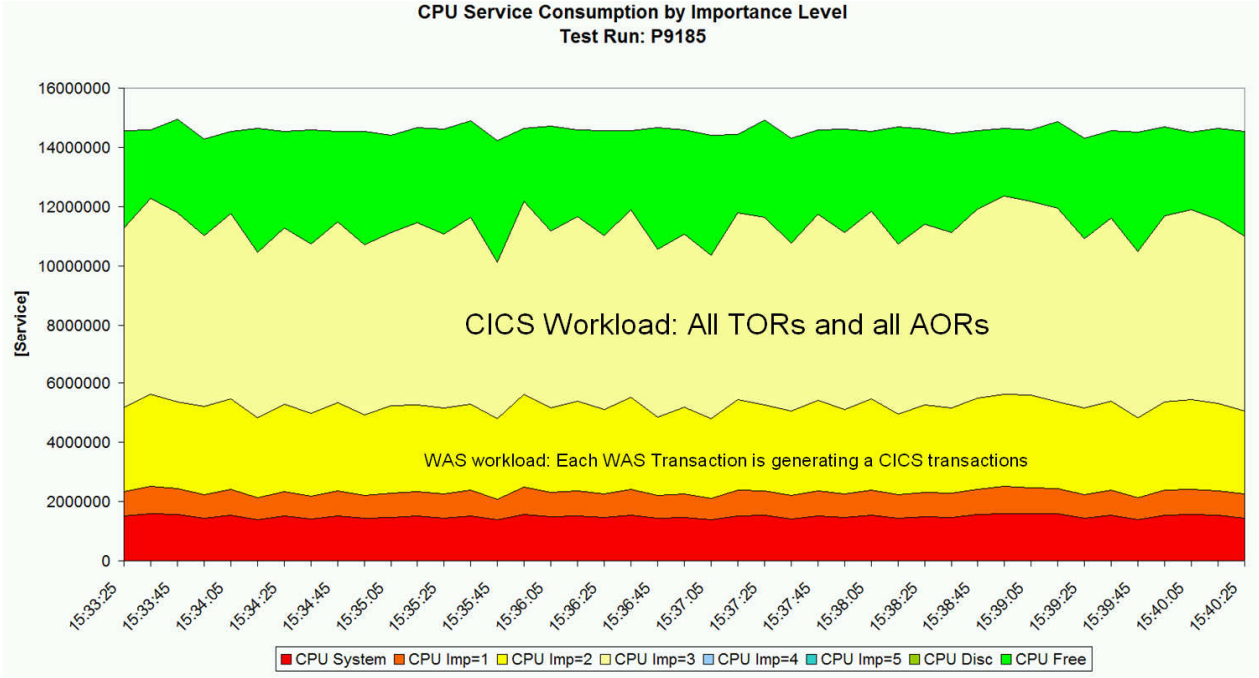
- z/Enterprise EC12 GA2 Support
- New Classification Qualifiers and Groups
- I/O Priority Groups
- Other z/OS V2.1 Enhancements
- Manage CICS Regions Using Goals Of: “BOTH”
- Service Stream Enhancements
- WLM Managed DB2 Bufferpools

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

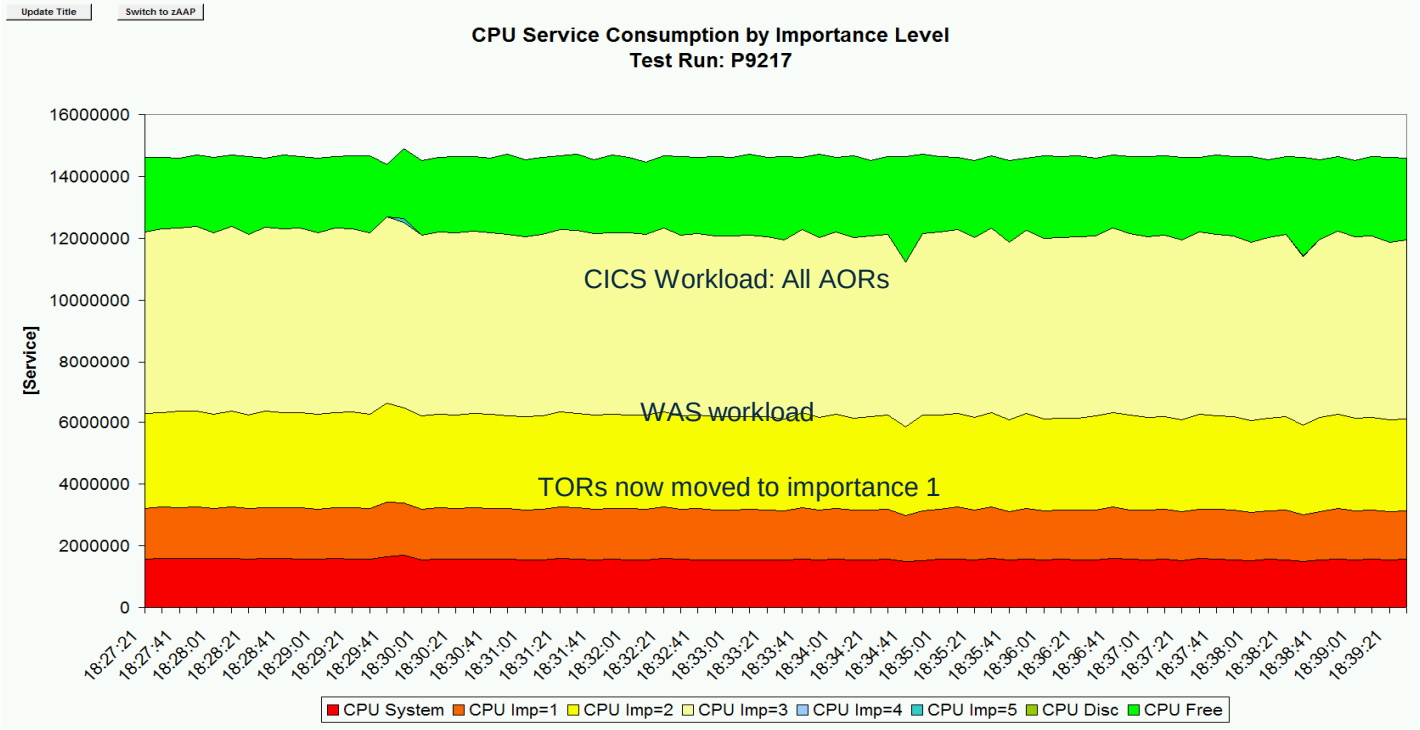
- **Example:**
  Workload: Websphere → CICS → DB2
  - Websphere receives work, sends it to CICS TORs which send it to AORs which execute DB2 calls
  - Classification: Websphere Imp=2 and all CICS Imp=3, managed towards response time goals
- ***Symptoms:***
  - Low system throughput
  - Relatively high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

Manage Regions Using Goals Of: “BOTH”



- 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.

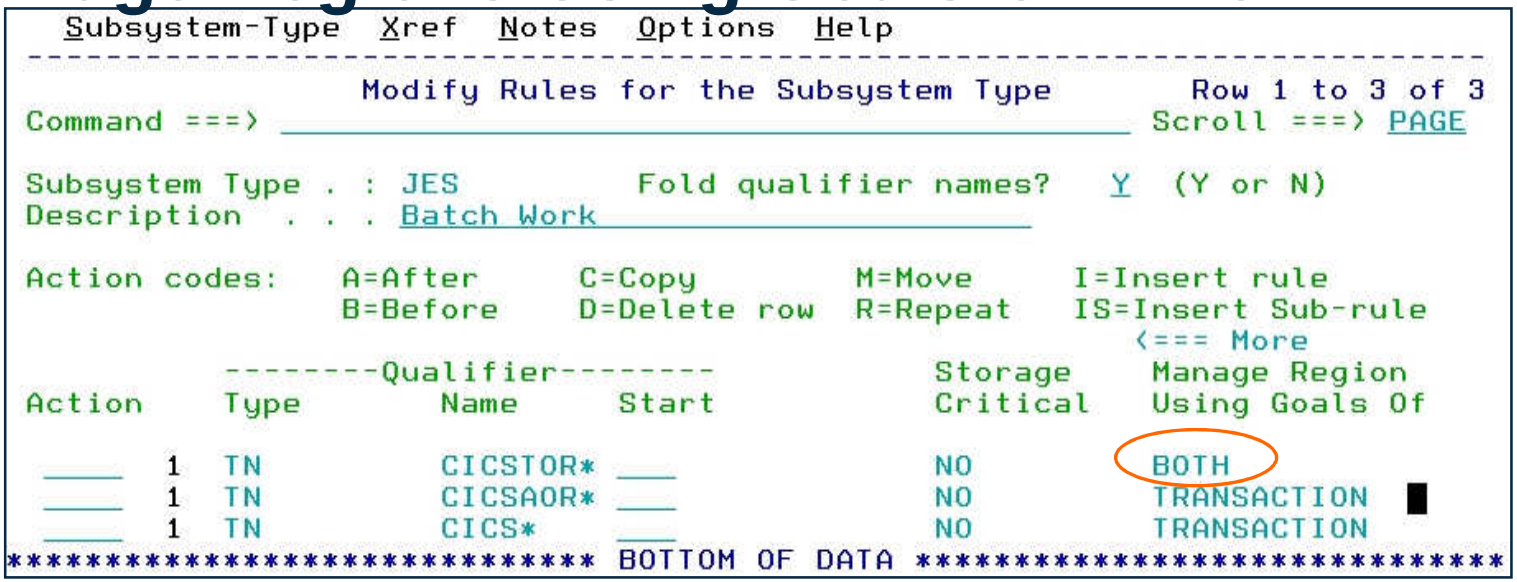
Manage Regions Using Goals Of: “BOTH”

- **Two possible alternatives**:
  (Move TORs to a service class with higher importance than AORs)
  - Option 1: Exempt **all** regions from being managed to 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”

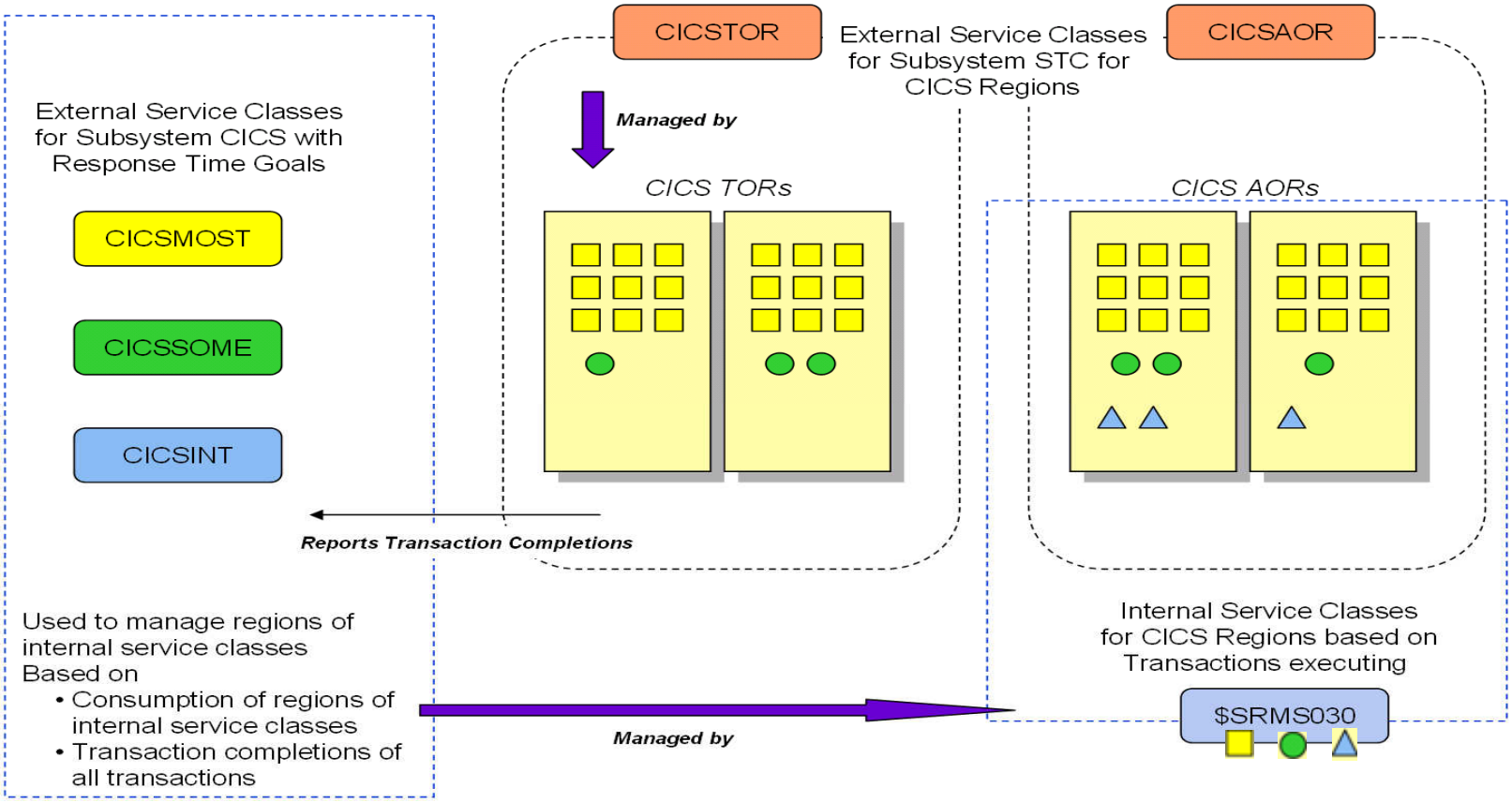


- Define STC service class for TORs which has a higher importance than the CICS service class with response time goals for the CICS work and AORs
- TOR: “Manage Regions by Goals Of: BOTH” option in WLM service definition
- AOR: “Manage Regions by Goals Of: TRANSACTION” (the default)

- Result:
  - WLM will manage the TORs towards the goals of the STC service class
  - And 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

Manage Regions Using Goals Of: "BOTH"

Structure of Service Classes



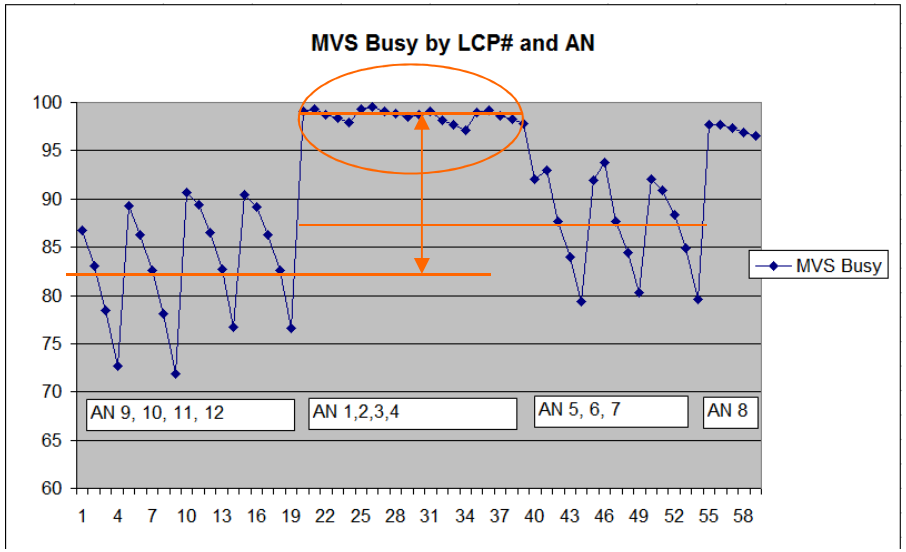
- TORs are now managed towards the goal of the service class CICSTOR
  - They still report their transaction completions for management
- AORs are still managed towards the goals of the CICS transaction service classes
- CICSTOR should be defined at a higher importance than the CICS service classes

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”
- Service Stream Enhancements
- WLM Managed DB2 Bufferpools

OA42185 / OA44337: 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%, while other nodes have plenty of free capacity.
  - Can typically be seen only with very stable workloads utilizing ≥ 30 CPs.



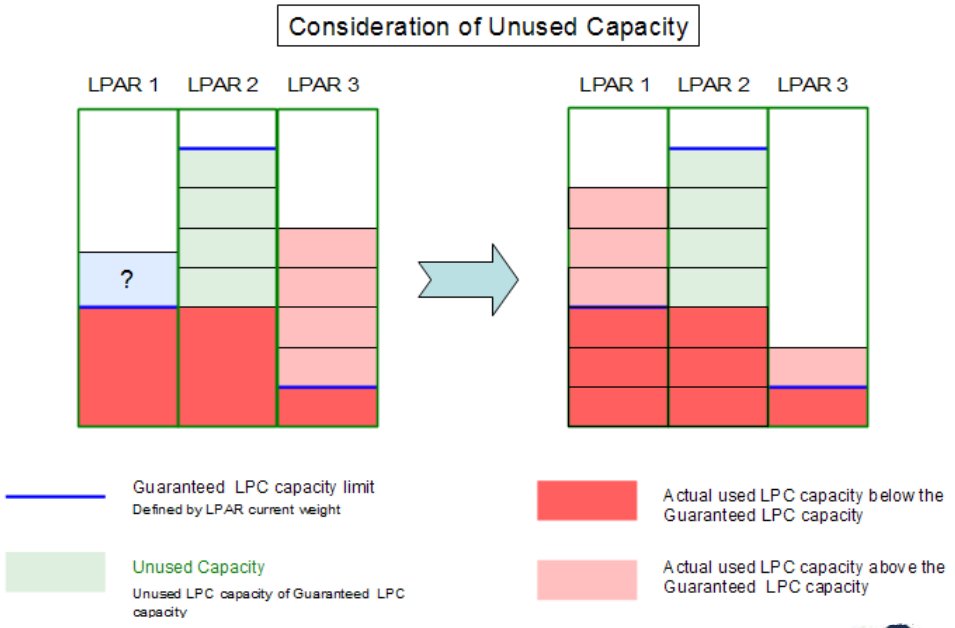
- The WLM algorithms were changed to more aggressively balance the work unit as the number of processors grow - resulting in a more even distribution of work across nodes.
  - Note: It is perfectly fine to see uneven processor utilization within nodes, for I/O enabled processors and also some uneven utilization across nodes (without overloading nodes, though)
  - z/OS V2.1: OA42185; z/OS V1.13: OA44337

Service Stream 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



Improvements for DB2 stored procedures and IDAA environments

- DB2 PM90151
  - In the case where a stored procedure spawns a thread and the spawned thread calls another stored procedure, the inner stored procedure can exceed the STORTIME zparm.
    - Can occur if WLM needs to start an additional server to satisfy the request.
  - With this APAR change, DB2 will use the DEPENDENT(YES) attribute when inserting the WLM request to schedule the stored procedure
  - Provided there are system resources available, WLM will give increased priority to this request. This should help prevent the sqlcode471 rc00E79002.
- WLM OA43538 (z/OS V1.12, z/OS V1.13, V2.1): “Unbound Servers”
  - Server address spaces, such as for DB2 Application Environments were not started due to incorrect assessment of available capacity
    - Symptom could be DB2 stored procedure timeouts with SQLCODE -471
  - Could occur even when minimum number of servers were requested via MNSPAS=n parameter

Service Stream Enhancements for more aggressive Blocked Workload support (OA44526)

- Problem addressed:
  - The current minimum value that can be specified for the Blocked Workload interval threshold BLWLINTHD is 5 sec.
    DB2 could profit from earlier or more frequent trickling.
- More aggressive specifications will be enabled by OA44526
  - New lower limit is 1 sec
  - BLWLINTHD default and BLWLTRPCT remain unchanged
    - Consider lowering BLWLTRPCT with very small BLWLINTHD values if amount of trickle cycles that may be handed out is a concern.

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”
- Service Stream Enhancements
- WLM Managed DB2 Bufferpools

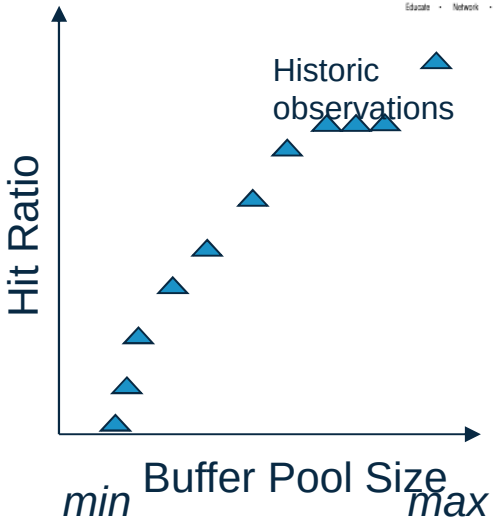
WLM-Managed bufferpool (AUTOSIZE BPs) update for z/OS V2.1

- REDP-5092-00: “Buffer Pool Monitoring And Tuning” says:

  AUTOSIZE(YES) does not require z/OS 2.1, but IBM recommends that you not use it until you have migrated to z/OS 2.1. Still, no customers have tested this function yet in production with z/OS 2.1. If you want to use it, you need to test it before considering it for production.

- DSNB555I @DBX2 WLM RECOMMENDATION TO ADJUST SIZE FOR BUFFER POOL BP11 HAS COMPLETED OLD SIZE = 1065 NEW SIZE = 1130

Changes to WLM-Managed DB2 Bufferpool Adjustment in z/OS V2.1



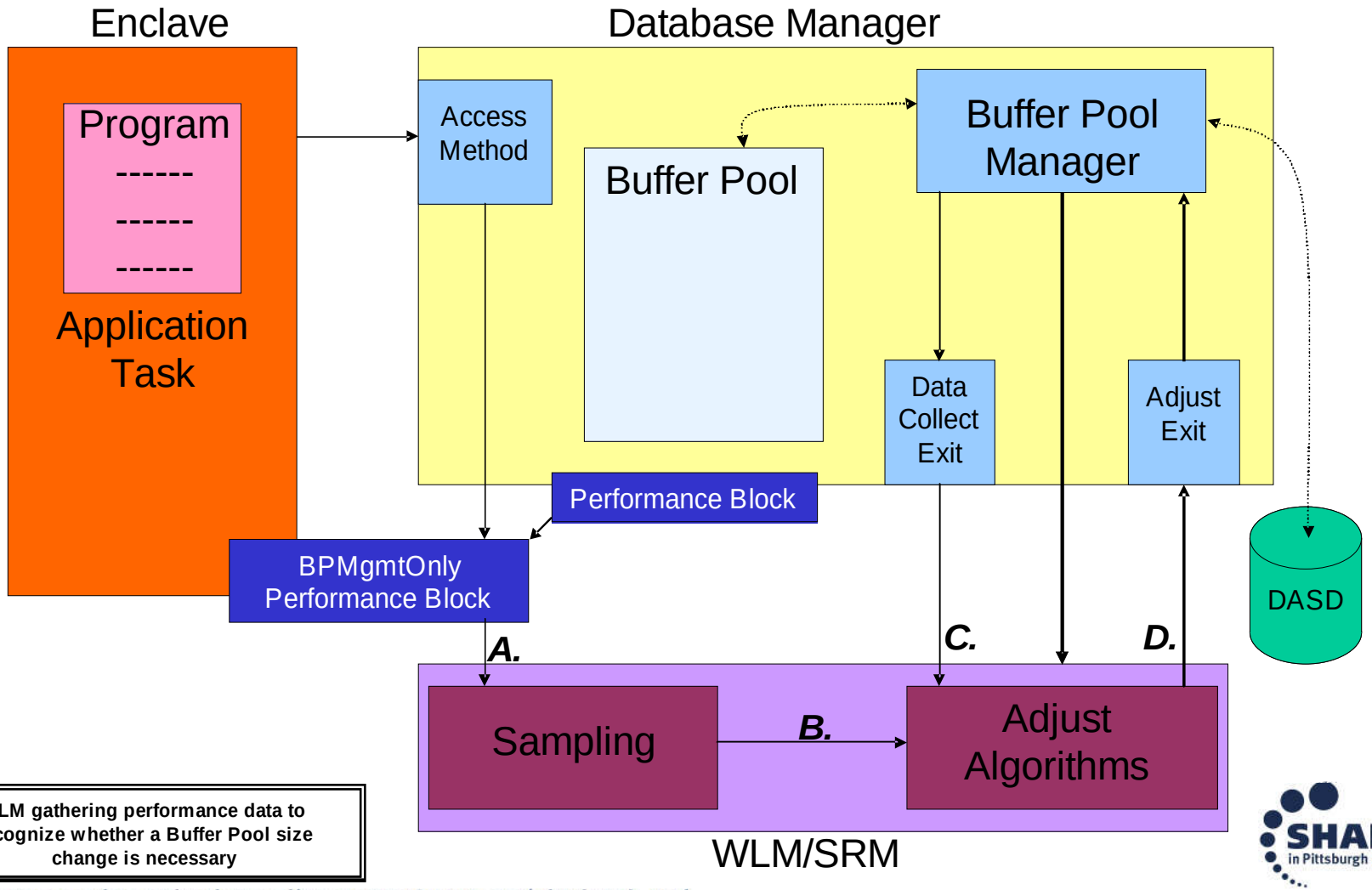
A bufferpool can be increased when

- Performance index impacted and buffer pool delays are significant contributor

A bufferpool may shrink…

- Due to donation to a suffering service class period
  - May suffer storage related delays
- Due to regular housekeeping cycles
  - Select one BP per 10 sec interval:
    - BP idle - had no references
    - No delays, i.e. 100% hit ratio
    - Is there any period of which the PI won't be impacted if buffer pool size reduced?
    - Any bufferpool may change no more than once a minute

Overall Flow WLM-managed DB2 Bufferpools



WLM gathering performance data to recognize whether a Buffer Pool size change is necessary

z/OS Workload Management - More Information -

- z/OS WLM Homepage:

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

- z/OS MVS documentation
  - z/OS MVS Planning: Workload Management: 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 publications*:
  - 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