z/OS Workload Manager

z/OS Release 8 Update

Horst Sinram
WLM Development
IBM Corporation, Boeblingen, Germany
mailto:sinram@de.ibm.com

9/30/2006

ON DEMAND BUSINESS™ = Make it happen now

Trademarks

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

APPN*	HiperSockets	OS/390*	VM/ESA*
CICS*	HyperSwap	Parallel Sysplex*	VSE/ESA
DB2*	IBM*	PR/SM	VTAM*
DB2 Connect	IBM eServer	Processor Resource/Systems Manager	WebSphere*
DirMaint	IBM e(logo)server*	RACF*	z/Architecture
e-business logo*	IBM logo*	Resource Link	z/OS*
ECKD	IMS	RMF	z/VM*
Enterprise Storage Server*	Language Environment*	S/390*	z/VSE
ESCON*	MQSeries*	Sysplex Timer*	zSeries*
FICON*	Multiprise*	System z	z9
GDPS*	NetView*	System z9	
Geographically Dispersed Parallel Sysplex	On demand business logo	TotalStorage*	
		Virtualization Engine	

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries

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

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

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

* 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.

ENABLING BUSINESS.
A THROUGH Z.

ON DEMAND BUSINESS™ = Make it happen now

Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ Preview: Group Capacity Limit

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

ON DEMAND BUSINESS™ = make it happen now

Overview of New Functions (1) – zIIP and zAAP

Function \ Release	z/OS R8	z/OS R7 w/ JBB772S	z/OS R7	z/OS R6 w/ JBB77S9	z/OS R6
zIIP Reporting Support	☺	☺ +OA16005		☺ +OA16005	
zAAP Management Support	☺				
zAAP Processing Enhancements	☺	☺	OA14131 + OA13953	☺	OA14131 + OA13953
zAAP Reporting Support	☺				
z/OS R8 Coexistence Support	n/a	OA13837	OA13837	☺	OA13837

ON DEMAND BUSINESS™ = make it happen now

Overview of New Functions (2)

Release / Function	z/OS R8	z/OS R7	z/OS R6	z/OS R5	z/OS R4
z/OS R8 Coexistence	n/a	OA13837			
RSM/SRM Support for >128GB real storage	☺				
IEAOPT MCCAFCTH Enhancement	☺	OA14409			
WLM Support of Multiple Subchannel Sets (z9-109)	☺	☺			
Enhancements in LPAR free capacity computations	☺	☺	OA10006		
Routing (Storm Drain) Enhancements	☺	OA14310 (planned)			
Sysplex Routing Enhancements	☺	☺			
Dynamically handle processor speed changes	☺	☺ OA12155	OA07510 OA12155	OA07510	

ON DEMAND BUSINESS™ = make it happen now

Overview of New Functions (3)

Enhancement	z/OS R8	z/OS R7	z/OS R6	z/OS R5	z/OS R4
Process Entitlement Type 2 and 3 Resource Groups	☺	z/OS R8 Coexistence Support (OA13837) on previous chart			
EWLM Service Class Correlation	☺	OA12784			
Lightweight EWLM Instrumentation for DDF	☺	OA12005			
EWLM Monitoring Support 1	☺		OA07196		
WLM-managed Initiator Start Considers System Affinities	☺	OA10814			
Controlling Initiator Dispatch Priority in Goal Mode	☺				OW55344

ON DEMAND BUSINESS = make it happen now

Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ Preview: Group Capacity Limit

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

zSeries Application Assist Processor (zAAP)

§ **What is it?**

– Powerful z/OS Java execution environment

– Allows installations to offload Java work from general purpose CPs to zAAPs

§ **Advantages**

– Significantly less expensive than general purpose CPs

– Possible reduction of demand on general purpose CPs

– Upgrade of additional zAAP processing power does not affect total MSU rating or machine model designation

– No IBM software charges on zAAP capacity

IBM eServer zSeries 990, 890, System z9-109 or later

ON DEMAND BUSINESS = make it happen now

zAAP Exploitation

§ **zSeries Application Assist Processors (zAAP) can only be deployed on IBM eServer zSeries 890 or 990 or later with z/OS V1.6 or later and the IBM SDK for z/OS Java 2 Technology Edition, V1.4 with PTF.**

§ **The following subsystems provide zAAP execution potential:**

– **WAS 5.1**
– IMS V7, V8, V9
– CICS 2.3
– DB2 V7, V8
– WBI Brokers V5
– **Java batch workload**

§ **Additional information can be found on:**

– zAAP home: http://www.ibm.com/servers/eserver/zseries/zaap/
– FAQ: http://www.ibm.com/servers/eserver/zseries/zaap/faq/
– Redbook: http://www.redbooks.ibm.com/redbooks/pdfs/sg246386.pdf
– Whitepaper: http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10337

ON DEMAND BUSINESS™ = make it happen now

Initial Handling of zAAP-Related IEAOPTxx Parameters

IFACROSSOVER	IFAHONORPRIORITY	Meaning
<u>YES</u>	<u>YES</u>	§ **CP's help zAAP's honoring priorities until discretionary work is reached.** § **CP's always help zAAP discretionary work after all CP work.** – With soft-capping in effect: See next row
YES	NO	§ **CPs always help zAAP's after all CP work**
NO	YES	
NO	NO	§ **CPs never help zAAPs, except when no zAAPs are operational**

Changed

ON DEMAND BUSINESS™ = Make it happen now

OA14131/OA13953: IFAHONORPRIORITY Enhancements

IEAOPT Parameters		New Behavior
IFACROSSOVER	**IFAHONORPRIORITY**	R6 and R7 with OA14131+OA15297/OA13953
YES	**YES**	Standard processors can run zAAP eligible work in priority order **if the zAAPs become unable to process all queued work** Notice**: regular CPs are asked for help therefore not all CPs may process zAAP work at the same time**
NO	**YES**	
YES	**NO**	**unchanged**
NO	**NO**	

§ **Old Behavior**
 – IFAHONORPRIORITY became only effective if IFACROSSOVER was set to YES

§ **New Behavior** (R6 and R7 with OA14131+OA15297/OA13953)
 – This change is intended to allow more zAAP eligible work to run on zAAP processors while still remaining responsive to the zAAP demand
 – IFAHONORPRIORITY is independent from IFACROSSOVER
 – IFAHONORPRIORITY now has the meaning that the zAAPs may "**need help**"

ON DEMAND BUSINESS™ = make it happen now

WLM Treatment of the zAAP Resource

§ **A zSeries Application Assist Processor is a new resource type that WLM is aware of**

  – Contributes using and delay samples

  – Contributes service times

  – zAAP utilization is reported by RMF (SPE OA05731)

§ **Up to z/OS 1.7 zAAPs are managed by WLM as extension of CPs**

  – Java work executing on zAAPs inherits the dispatch priority from its execution on regular CPs

  – Execution is accounted for in execution velocity and goal achievement (PI)

§ **zAAP management support**

  – Beginning with z/OS 1.8 work on zAAP is managed independently

§ **Not included in routing decisions**

§ **zAAP service not included in defined capacity computations and resource group management**

§ **zAAPs are not varied by IRD Vary CPU Management**

§ **On z9-109 specialty engines, such as zAAPs, are in separate pools.**

  – On z890 and z990: zAAPs inherit weight from CP pool but are part of the ICF/IFL pool

ON **DEMAND BUSINESS** = Make it happen now

Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ Preview: Group Capacity Limit

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

New IBM System z9 Integrated Information Processor (IBM zIIP)

§ **New specialty engine for the System z9 EC and BC mainframes designed to help:**

– Customers integrate data across the enterprise

– Improve resource optimization

§ **z/OS manages and directs work between the general purpose processor and the zIIP**

– Number of zIIPs per z9-109 not to exceed number of standard processors

– No changes anticipated to DB2 for z/OS V8 applications

§ **DB2 for z/OS V8 will be first IBM exploiter of the zIIP with**

– System z9 EC/BC

– z/OS 1.6 or later

– DB2 for z/OS V8

§ **More information available on the zIIP web site**

– http://www.ibm.com/systems/z/ziip/about.html

ON DEMAND BUSINESS™ = Make it happen now

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

IBM

How does the zIIP work .....

Work on z/OS may have all or a portion of its enclave Service Request Block (Enclave SRB) work directed to the zIIP. The types of DB2 V8 work listed below are those executing in enclave SRBs, portions of which can be sent to the zIIP.

Example 1 = Distributed SQL requests (DRDA)

Queries that access DB2 for z/OS V8 via DRDA over a TCP/IP connection are dispatched within z/OS in enclave SRBs. z/OS directs a portion of this work to the zIIP.

Example 2 = Complex parallel query (BI)

Complex star schema parallel queries will now use enclave SRBs. z/OS directs a portion of this work to the zIIP.

Example 3 = DB2 utilities for index maintenance

DB2 utilities LOAD, REORG, and REBUILD will now use enclave SRBs for the portion of the processing that is related to index maintenance. z/OS directs a portion of this work to the zIIP.

ON DEMAND BUSINESS™ = Make it happen now

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

IBM

Specialty engines

z/OS LPAR WAS Application

zAAP

TCP/IP HiperSockets

z/OS LPAR

DRDA

DB2

DRDA

Application not on System z

TCP/IP

WAS Application

Star Schema

zAAP

Application

CP

zIIP

DRDA

TCP/IP HiperSockets

Linux on System z LPAR Application

IFL

The IBM System z9 specialty engines can operate on the same machine together

(shown at left, instances where zIIP can be employed)

Handling of zIIP Work

§ **In many respects WLM/SRM support of zIIPs is equivalent to the zAAP monitoring support with the zAAP processing enhancements**

   – But: support only available via distinct FMIDs

§ **zIIP work managed as an extension of CP work**

§ **zIIP work will flow over to general purpose CPs**

   – Like zAAP work with the zAAP processing enhancements

     • "needs help"

   – But: No external controls

§ **Not included in routing decisions**

§ **zIIP service not included in defined capacity computations and resource group management**

§ **zIIPs are not varied by IRD Vary CPU Management**

§ **On System z9 specialty engines, such as zIIPs, are in separate processor pools.**

zIIP Reporting via RMF

§ WLM Services provide data on zIIP usage comparable to data on zAAP

§ **RMF Monitor I**

 – CPU and Workload Activity reports

§ **RMF Monitor III**

 – CPC, SYSINFO and ENCLAVE reports

§ **zIIP Usage projection available without real zIIPs in the configuration**

 – Via IEAOPT parameter PROJECTCPU=YES

§ **Detailed information available via http://www.ibm.com/systems/z/ziip/**

 – See "Introduction to System z9 Integrated Information Processor (zIIP)" at
 http://publibz.boulder.ibm.com/zoslib/pdf/zosr6r7ziip.pdf
 and the Program Directory
 http://publibz.boulder.ibm.com/epubs/pdf/i1007730.pdf

ON DEMAND BUSINESS = Make it happen now

Handling of zIIP Work

In this simple scenario we see that zIIP work is can be taken over smoothly by CPs.

The upper figure shows the service consumed on CPs (by importance level).
The lower figure shows the total service consumed by the service class that the zIIP work is using (by processor type).

Imp1 work is started on zIIPs such that work can be fully contained by the 2 zIIP engines.

In addition, Imp 3 work is started on CPs.

When one of the zIIPs is taken offline to simulate a shortage of zIIP resources, some of the work is taken over by CPs (orange area).

When the zIIP is becomes online again, the CPs no longer need to help the zIIPs.

The total service of the zIIP work's service class is decreased during the period of reduced zIIP capacity, because the share of the zIIP work running on CPs needs to compete against CP work.
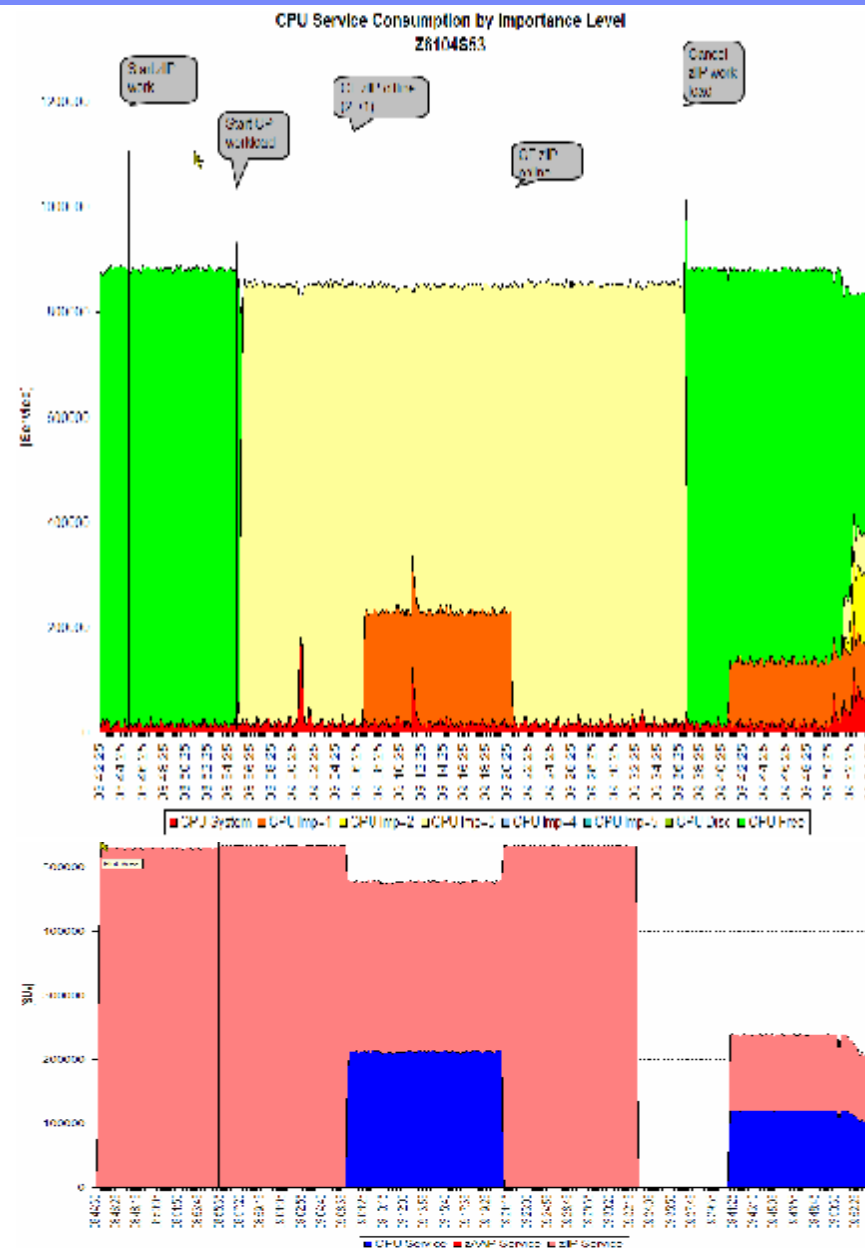
Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ Preview: Group Capacity Limit

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

ON DEMAND BUSINESS = make it happen now

RSM/SRM Support for Large Real storage

§ **Support up to 4 Terabytes of real storage**

§ **Page Replacement and UIC**

– The CPU cost of the current function continues to grow as we move to systems with more and more real storage while its effectiveness is diminishing. This algorithm was changed to perform more efficiently in large real systems.

§ **Physical Swap Processing**

– Physical Swaps to Auxiliary storage were eliminated

§ **Pageable Storage Shortages**

– Used to be resolved by physical swaps

– They are now resolved by frame exchanges

Page Replacement – pre z/OS 1.8

LSW / OUT

AS 3

IN

AS 1

AS 2

Page Datasets

§Today's LRU algorithm

§Implemented by keeping an Unreferenced Interval Count (UIC)

§ Available Frame Queue (AVQ) Low/Okay thresholds

§UIC update process runs periodically to update the UIC of each in-use frame

§Stealing starts when we go below AVQ Low and continues until AVQ Okay is reached

§Stealing happens on an address space basis

§Oldest frames are stolen first

ON DEMAND BUSINESS™ = Make it happen now

Page Replacement – z/OS 1.8

§Why Change?

§Disruption of the UIC update process is intolerable as the amount of real storage allocated to address spaces increases

§The new page replacement algorithm

§Runs when the Available Frame Queue needs to be replenished

§Stealing happens on a global basis

Page Datasets

make it happen now
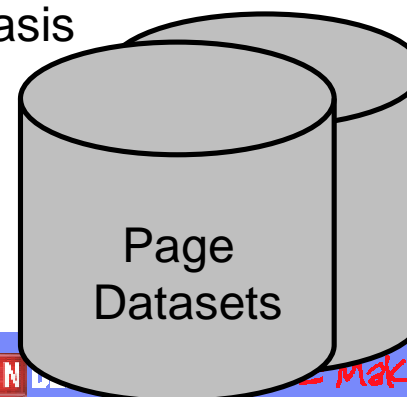
UIC Calculation in pre z/OS 1.8

LSW / OUT

AS 3

IN

AS 1

AS 2
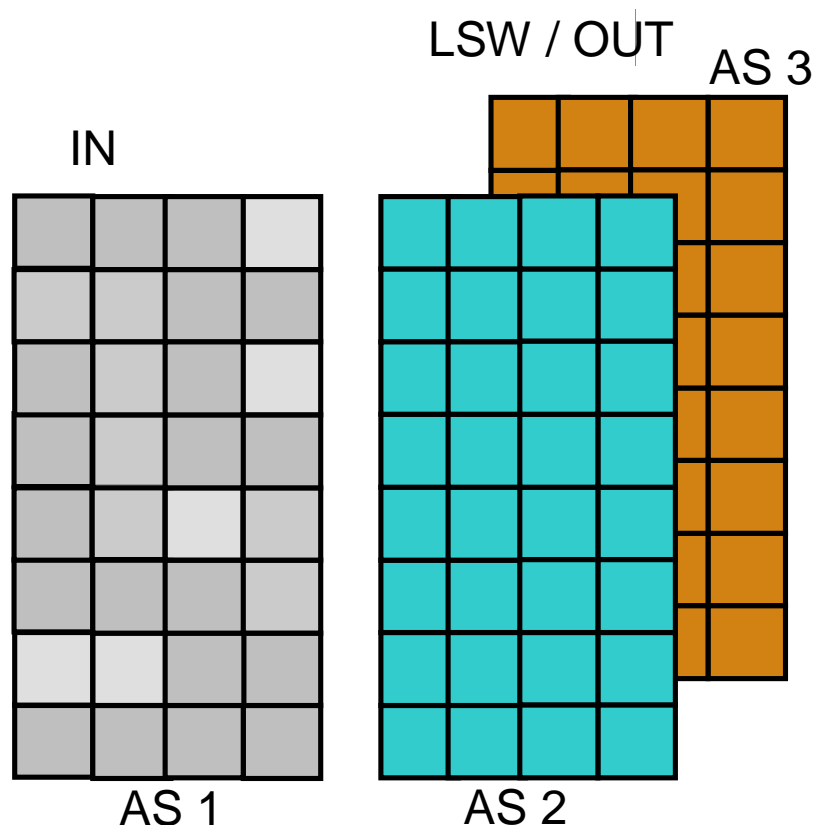
§The UIC Update process is scheduled to run every 10 seconds

§This process analyzes every pageable frame from each address space on the Swapped-IN queue

§The analyzed frames are counted and dependent on the age distributed into buckets.

§The oldest frame in an address space represents the address space High UIC

§The highest address space UIC represents the Highest System UIC (MCVSTCRI)

UIC Calculation in z/OS 1.8 …
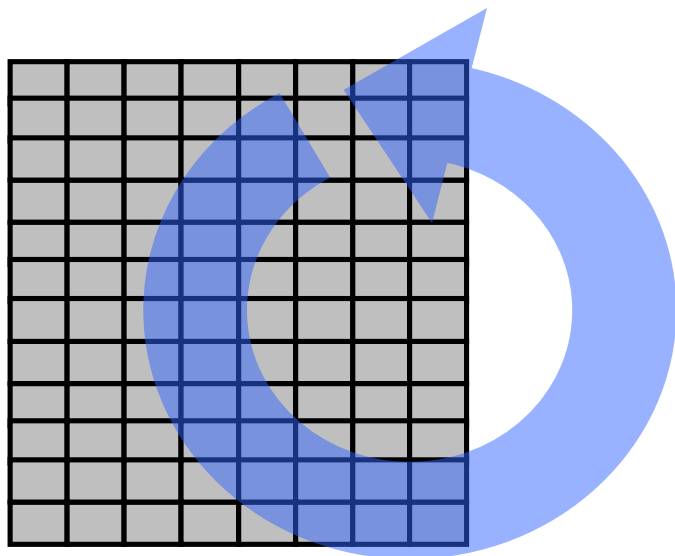
§There are three different UIC's which can
be displayed by the performance
monitors:
  §Current UIC
  §Minimum UIC
  §Maximum UIC

§The page replacement algorithm that z/OS
uses was enhanced to more efficiently process
large amounts of real storage.

§z/OS 1.8 defines the UIC as a single walk
though the whole storage in seconds.

§As a result the UIC values you might see in an
RMF report will vary from 0 – 65535 (18 hours),
in the past the UIC value range was 0 - 2540.

  §The higher the UIC value the less
  contention for storage in the system

  §The lower the UIC value the more
  contention for storage in the system. A very
  low UIC indicates that the system is storage
  constrained.

ON DEMAND BUSINESS™ = Make it happen now

Physical Swap – z/OS 1.8 ...

§z/OS 1.8 will no longer physically swap address spaces to AUX

§ As the amount of fixed storage an address space owns increases, swapping an address space out to AUX is no longer viable.

§Analyze your pre-z/OS 1.8 system and pay special attention to the amount of physical swaps that occur on your system.

§Use the RMF CPU Activity Report to find out how many physical swaps occur in your particular system configuration (OUT Ready and OUT Wait).

§Next for each address space on the OUT Ready or OUT Wait queue you need to find out the number of fixed frames each address space owns.

```
                            C P U   A C T I V I T Y   R E P O R T

SYSTEM ADDRESS SPACE ANALYSIS              SAMPLES =  3,600

NUMBER OF ADDRESS SPACES

      ------------------- QUEUE TYPES -------------------    --------- ADDRESS SPACE TYPES ---------
      IN                OUT        OUT    LOGICAL   LOGICAL   BATCH      STC     TSO    ASCH     OMVS
      READY       IN   READY      WAIT    OUT RDY   OUT WAIT

MIN      1        54       0         2          0        83       5      119      10       0        3
MAX     10        85       0         0          0       118      19      154      12       0       23
AVG    2.4      56.4     0.0       1.2        0.0     105.4     6.1    141.7    10.5     0.0      3.5
```

ON DEMAND BUSINESS™ = Make it happen now

New Message to Indicate High AUX Usage

§New External Message

   §**IRA205I 50% AUXILIARY STORAGE ALLOCATED**
   The message is issued when more than 50% of the Auxiliary storage slots are in use.

   §This allows an automation product to add additional page datasets before a critical shortage occurs

   §The message is repeated every 2 hours as long as the Auxiliary slot usage is above 50%

ON DEMAND BUSINESS™ = Make it happen now

Migration & Coexistence Considerations RSM/WLM

§Analyze your pre-z/OS 1.8 system and pay special attention to the amount of physical swaps that occur on your system. In most systems physical swaps never occur, so the fact that physical swaps were eliminated in z/OS 1.8 should have no impact.

§Use the RMF CPU Activity Report to find out how many physical swaps occur in your particular system configuration (OUT Ready and OUT Wait).

§ If the RMF report shows many address spaces on the OUT Ready or OUT Wait queue, you should consider increasing the amount of real storage on your system.

```
                          C P U   A C T I V I T Y   R E P O R T

SYSTEM ADDRESS SPACE ANALYSIS                SAMPLES =  3,600

NUMBER OF ADDRESS SPACES

        ------------------- QUEUE  TYPES -------------------    --------- ADDRESS SPACE TYPES ---------
        IN                OUT       OUT     LOGICAL   LOGICAL    BATCH     STC     TSO     ASCH    OMVS
        READY       IN    READY     WAIT    OUT RDY   OUT WAIT

MIN        1        54      0         0        0         83        5      119      10       0       3
MAX       10        85      0         0        0        118       19      154      12       0      23
AVG      2.4      56.4    0.0       0.0      0.0      105.4      6.1    141.7    10.5     0.0     3.5
```

ON DEMAND BUSINESS™ = Make it happen now

OA14409 IEAOPT MCCAFCTH Enhancement

§MCCAFCTH (lowvalue, okvalue)   Available on z/OS R6 and above

§ Specifies the LOW and OK threshold values for central storage.

§The low value indicates the number of frames on the available frame queue when stealing begins.

§The okvalue indicates the number of frames on the available frame queue when stealing ends.

§SRM will automatically adjust the actual threshold values based on measurements of central storage usage

§ The *Initial* lowvalue and okvalue are calculated by the following formula:

§RCEAFCLO=MAX(OPT Parameter (lowvalue),
    400,
    0.2% of the pageable storage)

§RCEAFCOK=MAX(OPT Parameter (okvalue),
    600,
    0.4% of the pageable storage)

§But capped to 35/70k, respectively

§ The formulas above will enforce the value, if the OPT parameter has specified a to smaller value

§ RCEAFCLO will vary between Initial Value and 70K frames

§ RCEAFCOK will vary between Initial Value and 140K frames

Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ Preview: Group Capacity Limit

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

Preview: Group capacity limit on IBM System z9 EC and z9 BC

§ **IBM plans to make it possible to define a logical partition (LPAR) group capacity limit on System z9 servers.
This function will be designed to allow you to specify a group of LPARs and apply a capacity limit to this group. This is expected to allow the system to manage the group in such a way that the capacity used by the LPARs will not exceed the group limit.**

§ **When available, support of group capacity limit will be exclusive to System z9 EC and z9 BC**

ON DEMAND BUSINESS™ = Make it happen now

Group Capacity Limit Definitions

§ **LPAR Definitions**

– Defines a capacity group by entering the group name and group limit value for the partitions which should belong to the same group

- Group Limit and Group Name will be new input values on the HMC/SE
- Group Limit definition is independent from defined capacity definition
  - Both limits can be defined and both work together

§ **Capacity Group**

– Consists of multiple LPARs on the same CPC

– LPARs must run z/OS 1.8

- If they don't, the group limit may not be enforced correctly

– It is possible to define multiple groups on a CPC

– A partition can only belong to one group

– A capacity group is independent of a sysplex and an LPAR cluster

– WLM will only manage partitions with shared CPs

  - Dedicated partitions and partitions with wait completion equal to YES are ignored
    - > if they have been defined to a group they will be excluded from the group and only the partitions with wait completion equals NO and shared processors are managed towards the capacity limit

ON DEMAND BUSINESS™ = Make it happen now

Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

JES2 Improved Batch Initiator Balancing

§ **Problem prior to z/OS R8:**

– In a WLM managed initiator environment JES2 will preferably start jobs on the *submitting* system

§ **Enhancement in R8 attempts to use *approximately* the same percentage of active WLM-managed initiators in each service class on each system.**

– May help balance batch workloads across a MAS configuration within a sysplex

– In the scenario on the next chart we submit one job per minute (for a total of 50 min) on SYS1.

- With R7, the blue lines (Active servers = initiators) differ on the two systems. E.g. SYS1 peaks at 52, SYS2 at 45.

- With R8, the lines are very similar and peak at 50.

ON DEMAND BUSINESS™ = Make it happen now

JES2 Improved Batch Initiator Balancing - Comparison

z/OS R7

z/OS R8
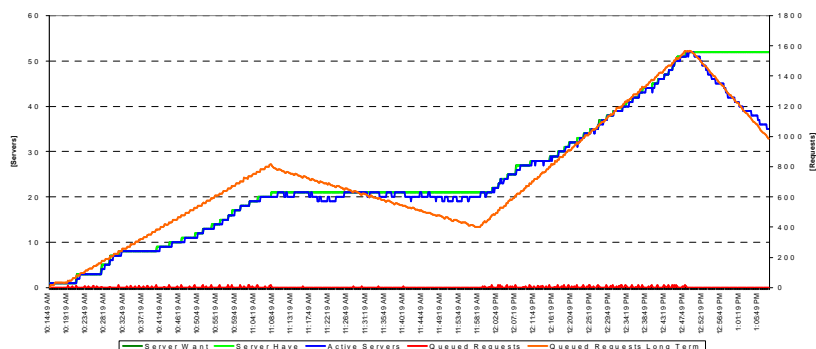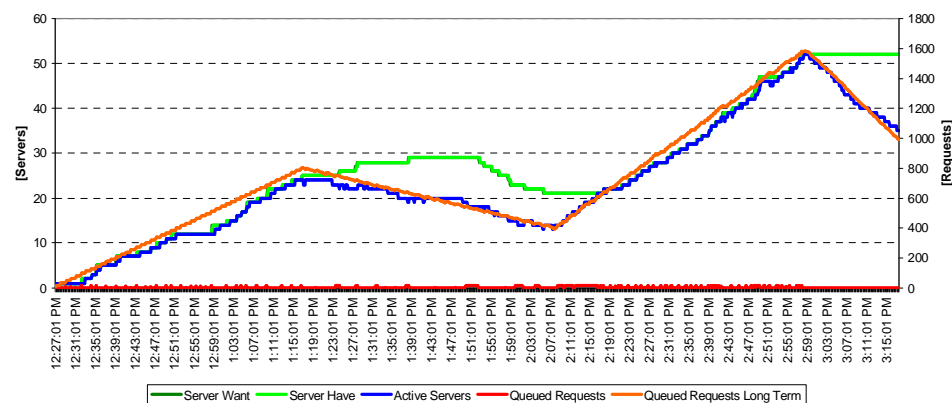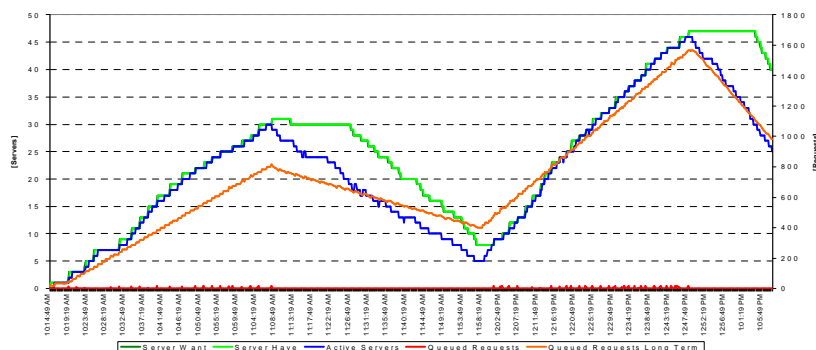
© 2006 IBM Corporation

ON DEMAND BUSINESS™ = make it happen now
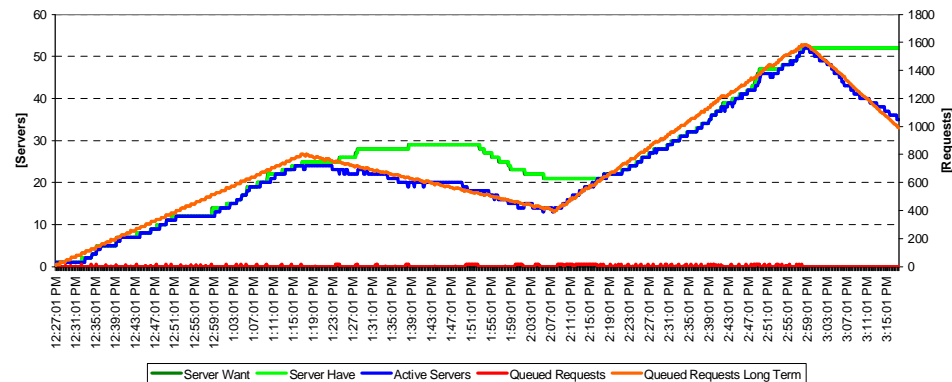
Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

Existing Resource Group Concept
Type 1 Resource Group

§ **Sysplex-wide defined in unweighted service units per second**

§ **Sysplex-wide managed**

§ **General Considerations**

– Multiple service classes may be assigned to a resource group

- With different utilizations on different systems

– Systems may have different capacities
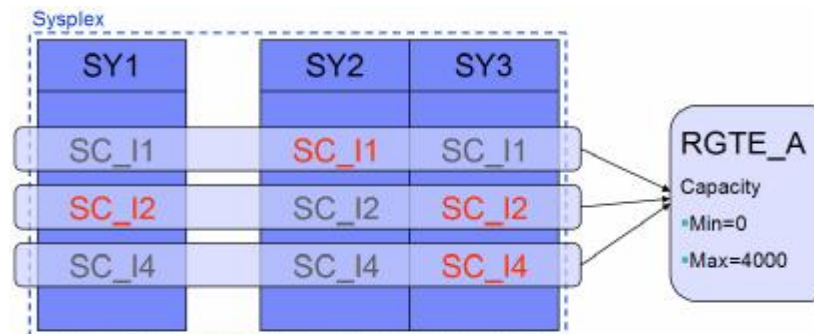
§ **Results:**

– Not easy to understand how much is consumed on which system

– Consumption depends highly on the capacity of the systems!

– Resource Group definitions need to be revisited every time

- Systems are upgraded
- Workload utilization changes

§ **Resource Groups are sometimes hard to understand**

Sysplex

SY1		SY2	SY3
SC_I1		SC_I1	SC_I1
SC_I2		SC_I2	SC_I2
SC_I4		SC_I4	SC_I4

RGTE_A
Capacity
- Min=0
- Max=4000

ON DEMAND BUSINESS™ = make it happen now

New Types of Resource Groups

§ **Sysplex-wide defined, but definition applies to each system**

§ **Managed by each system**

§ **General Considerations**

  – Multiple service classes can be assigned to a resource group but this has no sysplex-wide effect anymore

  – Definition is based on one of two possible units:

    • **LPAR capacity: based on system weight (+…)** ＥType 2 Resource Group
    • **LCP capacity** ＥType 3 Resource Group
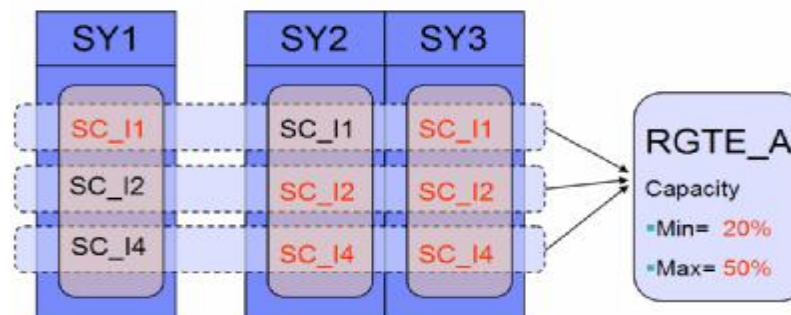
§ **Results**

  – New resource groups are managed by system thus they must be evaluated on a per system basis

  – Resource groups grow automatically if systems are upgraded

§ **Easier to understand**

Type 2 Resource Group Specification

§ **Capacity is specified as a percentage of the effective LPAR share**

- − Range 0..99

- − Effective LPAR share is the minimum of
  - LPAR share:
    CPC shared pool capacity · (LPAR weight / sum of weights)
  - Logical processor capacity
  - Defined capacity limit (only when soft capping is in effect)

- − Use WLM Planning: Appendix B - CPU Capacity Table
  to locate the CPC capacity that would be equivalent to your shared pool
  configuration

Numbers should be
meaningful for
configuration,
e.g. sum of all minimum
LPAR shares œ 99

$$RG_{[SUs]} = \begin{cases} Softcap_{[SUs]} \bullet RG_{Limit[\%]} & \text{if } softcap < ShareCapacity \text{ or } LCPCapacity \\ LCPCapacity_{[SUs]} \bullet RG_{Limit[\%]} & \text{if } LCPCapacity < ShareCapacity \text{ or } softcap \\ ShareCapacity_{[SUs]} \bullet RG_{Limit[\%]} & \text{if } ShareCapacity < LCPCapacity \text{ or } softca \end{cases}$$

$$ShareCapacity_{[SUs]} = CECCapacity_{[SUs]} \bullet LPARShare$$

$$CECCapacity_{[SUs]} = \text{Capacity based on shared physical processors for the CEC}$$

$$LPARShare = \frac{Weight(Current\ Partition)}{\sum_{i}^{all\ active\ partitions} Weight(i)}$$

$$LCPCapacity_{[SUs]} = \text{Capacity based on shared processors available to the LPAR}$$

Type 3 Resource Group Specification

§ **Capacity is expressed as an equivalent number of general purpose processors**

- – Scaled by 100, i.e. 100 equivalent to 1 CP
  Range: 0…999999

- – Use *WLM Planning: Appendix B - CPU Capacity Table*
  to locate the model that would be equivalent to your virtual processor configuration.

Numbers should be "meaningful" for configuration.

Comparison of Resource Group Types

Definition	Scope	Advantages	Possible Disadvantages
Type 1 RG: Service Units	**Sysplex**	§ **Allows balancing of resources across members in a sysplex**	§ **Requires adjustments for Migration** § **Difficult to monitor** § **Not applicable to constrained work on a single system in a sysplex environment**
Type 2 RG: % of LPAR	**System**	§ **Allows to control work on single members in a sysplex** þ **Stable for migrations**	§ **Perhaps difficult to understand what LPAR capacity really is**
Type 3 RG: % of LCPs	**System**	§ **Allows to control work on single members in a sysplex** þ **Easy and straight forward definition**	§ **Requires adjustments for migration**

ON DEMAND BUSINESS™ = make it happen now

Resource Groups: Administrative Application

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

Enter or change the following information:

Resource Group Name  . . . . : DIVB
Description  . . . . . . . . . Division B service

Define Capacity:
2   1.   In Service Units (Sysplex Scope)
    2.   As Percentage of the LPAR capacity (System Scope)
    3.   As a Number of CPs times 100 (System Scope)
Minimum Capacity . . . . . . 10
Maximum Capacity . . . . . . 20
```

§ **Resource group definition panel has been extended and does now allow for definitions of the three variations for resource groups**

1. In service units with a sysplex-wide scope (existing)

2. As a percentage of the LPAR capacity (new concept)

   - Based on LPAR weight
   - System-wide managed

3. As a percentage of the LCP capacity

   - System-wide managed

ON DEMAND BUSINESS™ = Make it happen now

Resource Groups: Example

Resource Group Overview
System: IRD1, Resource Group: ELPMAX , Type 2: 60% of LPAR

CEC Capacity	~ 111758 SU/s (2084-306)	
LPAR Capacity	~ 8697 SU/s Based on Weight (7.7%)	
ELPMAX Capacity	~ 5218 SU/s	60% of the LPAR capacity

Legend: Minimum Service Rate — Maximum Service Rate — Actual Service Rate — Capping Slices

ON DEMAND BUSINESS™ = Make it happen now
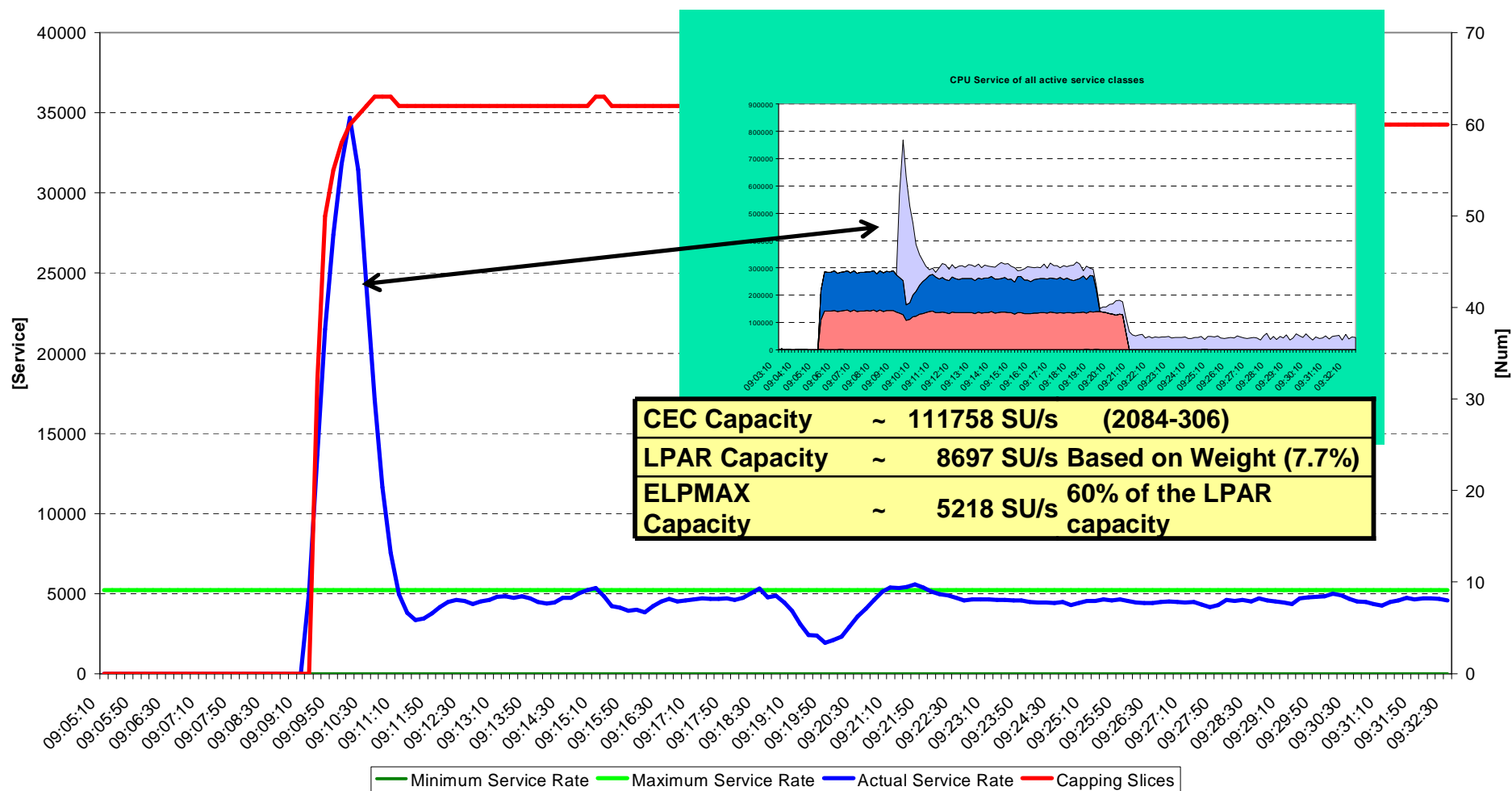
Resource Groups: Migration Considerations

§ **New system service classes and new resource group concept introduce incompatible changes to the WLM service definition**

 – Introduce new WLM policy level to LEVEL019

 – APAR OA13837 is required on pre-R8 systems

   • listed in GA22-7499 z/OS V1R8.0 Migration: z/OS V1Rx coexistence and fallback PTFs

§ **The compatibility APAR OA13837**

 – Ensures that type 2 and 3 resource groups are not managed on pre-1.8 systems

 – As long as policies are only installed / activated from back level systems the compatibility code is not required

   • Compatibility PTFs can be installed via rolling IPL

     – no Sysplex-wide IPL necessary

Resource Groups: WLM API Changes

§ **Resource Groups**

– IWMSVPOL Service

- The values for SVPOLGMN and SVPOLGMX (IWMSVPOL) mapping can now be in percentages
  - of LPAR capacity (SVPOLGPV bit turned on)
  - of single processor capacity (SVPOLGPC bit turned on)
- If both bits are off the values are in unweighted service units

– For mixed-release sysplexes (z/OS 1.8 and older versions)

- On pre-z/OS 1.8 releases the SVPOLMXS and SVPOLMNS bits indicate that NO limits have been specified.

Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

Routing Services: Before z/OS 1.7

§ **Step 1**

– Registration of servers by IWMSRSRG services

- Each server registers on the local system as a routing candidate
- WLM communicates registered servers and capacity data between systems
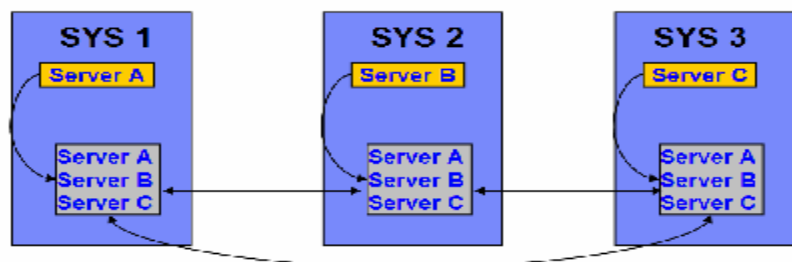
§ **Step 2**

– Router asks WLM for recommendations by using IWMSRSRS service

- WLM calculates a weight for each registered server based on capacity
  – Servers on systems with serious constraints are excluded (weight 0) unless all systems have serious constraints
- The returned weight reflects the available or replaceable system capacity

ON DEMAND BUSINESS = Make it happen now
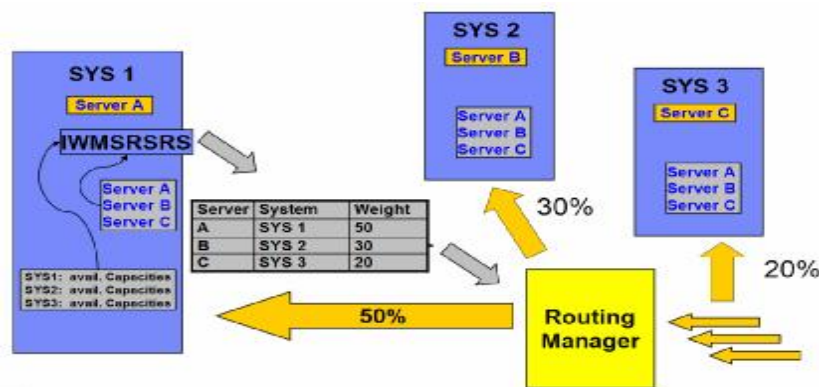
Sysplex Routing: Observations Before z/OS 1.7

§ **Current recommendation is purely capacity based**

– Weight reflects system capacity and is not server specific

– Some observed situations:

- If the LPAR is loaded with low important work, but the server has high importance
  → weight too low
- If the LPAR is lowly utilized, but the server has a bad performance index or is waiting for resources
  → weight too high

§ **TCP/IP stack registration does not reflect how work is processed on the system**

– TCP/IP address space is usually running in a high important service class

– The registered address space does not consume the work, it just passes the work to the real consumer on the system

§ **Abnormal conditions of the work receiver or consumer are not factored into the routing recommendations**

– This can lead to *storm drain* scenarios where a system seems to be eligible for processing many requests while in fact the work receiver or consumer is ill

– In many cases this requires that the work receiver consumer provides additional information to WLM

§ **LPAR Wait Time includes wall-clock time**

– In addition soft capping and LPAR weights are not correctly reflected

– This can result in an over estimation of the available capacity of a logical partition

– As a result to many requests may be routed to certain systems (partitions) in a sysplex while other systems receive to few requests

Sysplex Routing: Enhancements

§ **Current recommendation is purely capacity based**

- New option on routing services will allow a router to receive weights which consider the goal achievement of the work
- **z/OS 1.7   (IWMSRSRS FUNCTION=SPECIFIC)**
  - **May be exploited by DB2 in the future**

§ **TCP/IP only registers the stack and doesn't know how the work is being processed**

- Introduce new routing service e.g. for TCP/IP which allows to provide routing recommendations for TCP/IP backend applications
- **z/OS 1.7 (IWM4SRSC)**

§ **Abnormal conditions of the work receiver or consumer are not factored into the routing recommendations**

- In a first step the queue time of the work is accounted for DDF regions and its proportion to the execution time is factored into the routing recommendations
- **z/OS 1.7 (incorporated in IWMSRSRS FUNCTION=SPECIFIC)**

– Current activities to provide broader solutions

- Include abnormal terminations which are recorded by CICS
- New interface/option which allows a server to provide its state to WLM

§ **Measured Waittime is put in perspective with real available system and partition capacity**

- Over estimation is reduced
- This part is also available as **OA10006 on z/OS 1.4 and above**

Routing Services: Summary

Service	Function	Remark	Description
IWMSRSRS	**Existing function SELECT**		**Returns capacity for the system relative to other systems in the sysplex (accounts for the number of registered server instances)**
IWMSRSRS	**New parameter SPECIFIC**	**(1)**	**Returns capacity for the registered server relative to all other registered servers of the same type in the sysplex. Considers:** § **Goal Achievement (PI)** § **Queue Time for Enclaves** § Health Indicator
IWMSRSRG	New Parameter HEALTH		Allows the server which registers to provide a health indicator from ok=100 to not ok=0. The factor is considered as part of the weight. IWMSRSRG HEALTH can be updated by the server at any time
IWM4HLTH	New Service		For address spaces which are not registered and which want to set a health status. This status is factored into IWM4SRSC return data
IWM4SRSC	**New service** New parameter ABNORM_COUNT	**(1)**	**Returns capacity for another address space to which the request is provided by the registered server. Considers** § **Goal Achievement (PI)** § abnormal termination rate expressed as the number of abnormal terminations (as passed to WLM by the IWMRPT interface) per 1000 total terminations.

(1)   Goal achievement is derived from the service class the working is running in. Can be an enclave service class, a transaction manager service class or the service class to which the address space is classified too

ON DEMAND BUSINESS™ = Make it happen now

Sysplex Routing: Enhancements for IWMSRSRS Service

§ **New Function code „SPECIFIC"**

§ **Weight calculation: Product of four factors**

– System Utilization Factor:

   • Same as the resulting system weight for old „SELECT" function (as described in the earlier foils)

– PI Factor

   • This gives an indication of how good this server, respective the work that is related to this server, is achieving its goals as defined in the active WLM policy.

– Queue Time Ratio:

   • If the server owns independent enclaves, the ratio of queue time to elapsed time of those enclaves

– Health Indicator

   • A registered server can use the IWMSRSRG HEALTH= parameter to inform WLM about additional states which are unknown to WLM but should be factored into the returned weight

– Remark

   • If multiple servers are registered on the same system, the weight is divided by the number of the servers.

Sysplex Routing: New IWM4SRSC Service

§ **Characteristics**

– Provides routing recommendation for a given server

– System scope

– No registration of servers necessary

– The intention of this service is to obtain routing recommendations for work which is passed through the registered server (e.g. TCP/IP) to a server which consumes/processes the work

– Input: STOKEN that identifies the address space of a server

– Output: Recommendation of how good this server is suitable to receive work à the weight

  • A value between 1 and 64

§ **Weight calculation: Product of two (four) factors, scaled by 64**

– The PI Factor (Performance Indicator Factor)

  • This gives an indication of how good this server, respective the work that is related to this server, is achieving its goals as defined in the active WLM policy.

– The Importance factor

  • This is a measurement of how much CPU Capacity is displaceable by work of the server's importance, respective the work that is related to this server

– The health factor

  • This is an indicator an address space can set via the new IWM4HLTH service

– Abnormal termination rate

Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary

User Friendly Interface for WLM Administrative Application
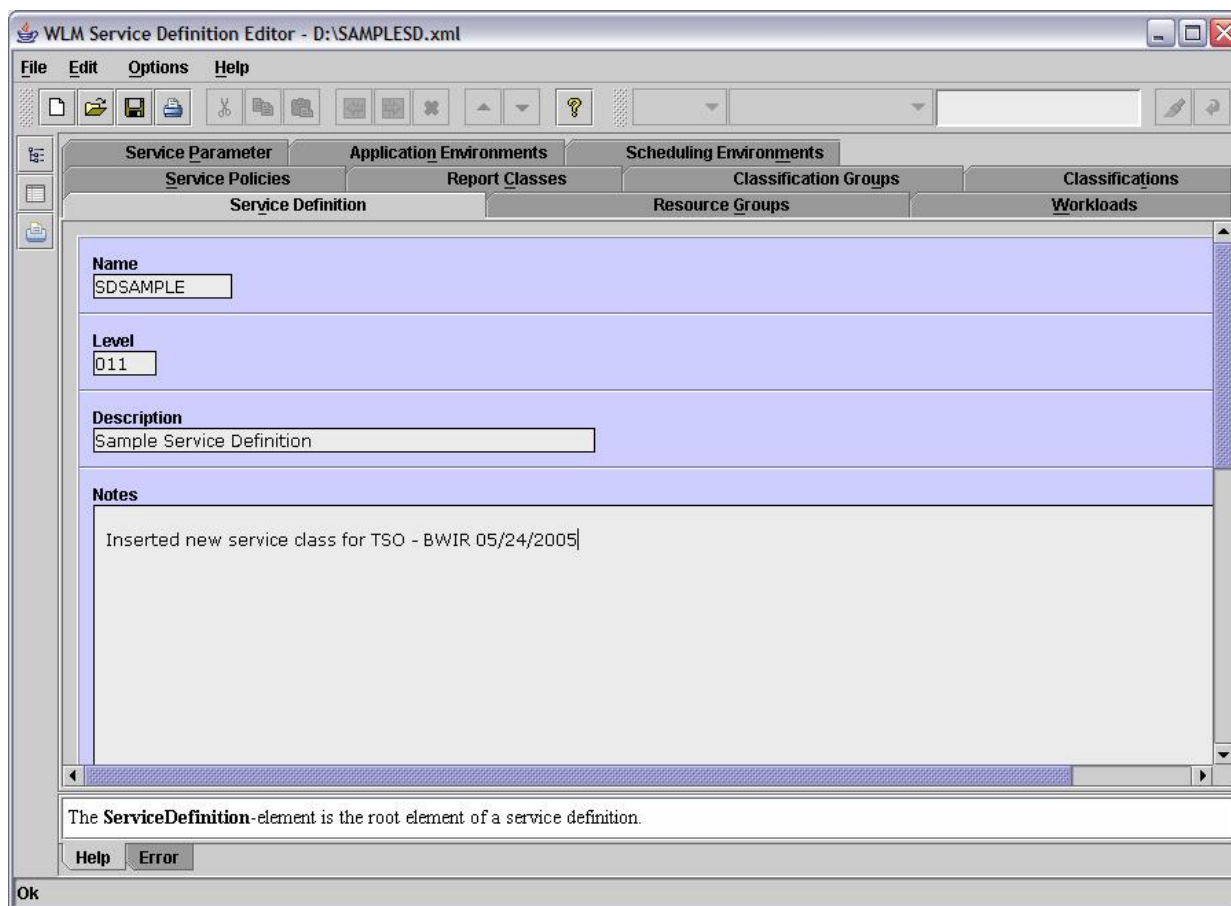
§ **Java application**

§ **Interface to WLM ISPF datasets**

§ **Different views to show relations between policy elements**

§ **Real-time error checking**

§ **Real-time help**

§ **No popup menus**

§ **Tabular input processing**

User Friendly Interface for WLM Administrative Application
Example: Workload and Service Class Panel/View

§ **Information is represented as tables**

§ **Direct manipulation of tables**

§ **Display of relations between policy elements**

§ **Real-time error checking**

§ **Context-sensitive help**

User Friendly Interface for WLM Administrative Application
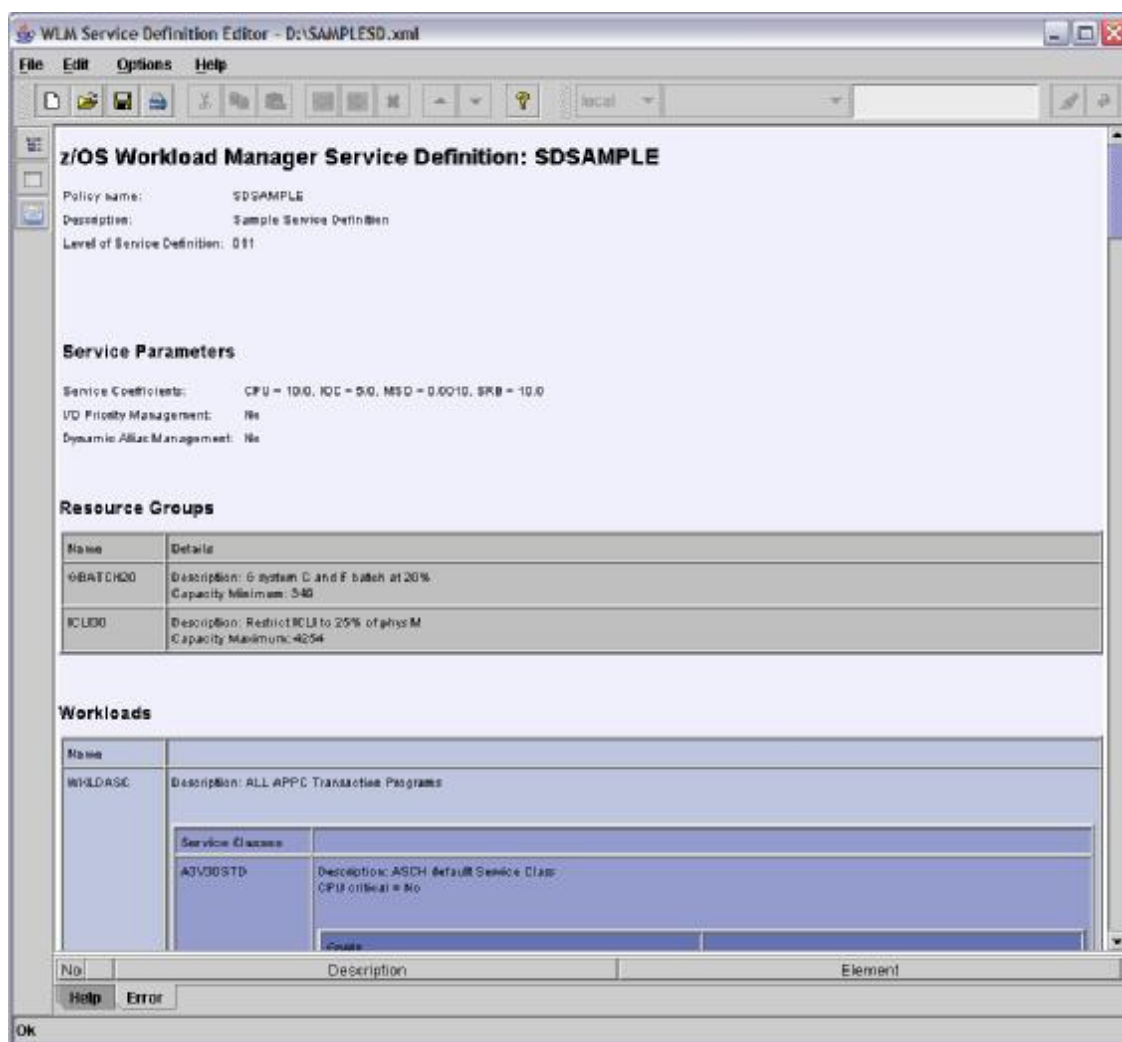## Example: Easy generation of new views

§ **Table columns can be moved as needed**

§ **Table rows can be sorted as needed**

§ **Rearranged tables can be edited**

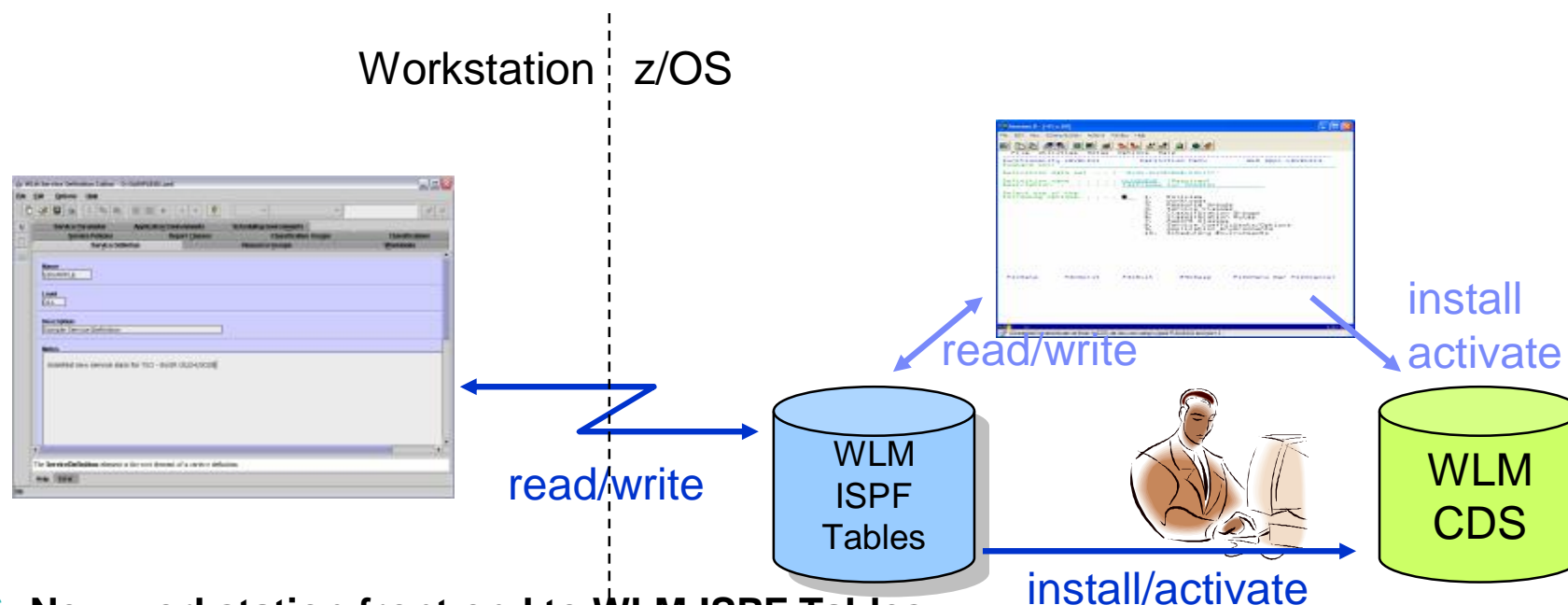§ **Search function for arbitrary text within tables**

WLM Administrative Application
Example: Print View

§ **Service Definition is converted to HTML-Document**

§ **In order to print HTML-Document the configured WWW-Browser of the Operating System is launched**

WLM Administrative Application
Link to existing WLM Administrative Application

Workstation ¦ z/OS

install
read/write
install
activate

read/write

WLM
ISPF
Tables

install/activate

WLM
CDS

§ **New workstation front-end to WLM ISPF Tables**

– FTP connection with automatic upload and download

– FTP connection profiles

§ **Policy install and activate via operator command**

Delivery:

– As a tool that can be downloaded via the WLM homepage
http://www.ibm.com/servers/eserver/zseries/zos/wlm/tools/

ON DEMAND BUSINESS™ = make it happen now

New ITSO Redbook

§ **System Programmer's Guide to: Workload Manager (SG24-6472)**

- Content:

    1. Introduction
    2. How WLM works
    3. WLM functions
    4. Implementation and workload classifications
    5. Batch considerations
    6. TSO, STC, and APPC workloads
    7. DB2 workload considerations
    8. WebSphere Application Server workload considerations
    9. UNIX System Services considerations
    10. Transactional workload considerations

- Download via:
    http://www.redbooks.ibm.com/abstracts/sg246472.html?Open

ON DEMAND BUSINESS™ = Make it happen now

Agenda

§ New function overview

§ zAAP Support

§ zIIP Support

§ RSM/SRM Support for Large Real storage

§ JES2 Enhancements for WLM-managed batch initiators

§ New Resource Group Types

§ Routing Enhancements

§ User Friendly Interface for WLM Administrative Application

§ Summary