



Session P03

WLM IRD and Defined Capacity

Ulrich Hild

IBM
SYSTEM z9 AND zSERIES EXPO
October 9 - 13, 2006

Orlando, FL



zSeries EXPO 2006 Orlando, FL

WLM Advanced Topics: IRD and Defined Capacity

Ulrich Hild
ulrich_hild@de.ibm.com
Session P03

zSeries® EXPO
FEATURING Z/OS, Z/VM, Z/VSE
AND LINUX ON ZSERIES

ON DEMAND BUSINESS™

Trademarks

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

AIX*	HiperSockets	Parallel Sysplex*
APPN*	HyperSwap	PR/SM
CICS*	IBM*	pSeries*
DB2*	IBM eServer	RACF*
DB2 Connect	IBM logo*	Resource Link
DFSMSdfp	IMS	S/390*
DFSMSdss	iSeries	S/390 Parallel Enterprise Server
DFSMSHsm	Language Environment*	Sysplex Timer*
DFSMSrmm	Lotus*	TotalStorage*
e-business logo*	Multiprise*	VTAM*
Enterprise Storage Server*	MVS	WebSphere*
ESCON*	NetView*	z/Architecture
FlashCopy*	Notes*	z/OS*
GDPS*	On demand business logo	z/VM*
	OS/390*	zSeries*

* 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 Linux 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.

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.

Agenda

§ Intelligent Resource Director

- CPU Management
- Dynamic Channel Path Management
- Channel Subsystem Priority Queuing

§ WLM Defined Capacity

Acronyms

- § **CP – (General purpose) processors**
- § **CPC – Central processor complex**
- § **DCM – Dynamic Channel Path Management**
- § **IRD – Intelligent Resource Director**
- § **LCP – Logical processors**
- § **LPC – LPAR cluster**
- § **PCP – Physical processors**

What you won't find here...

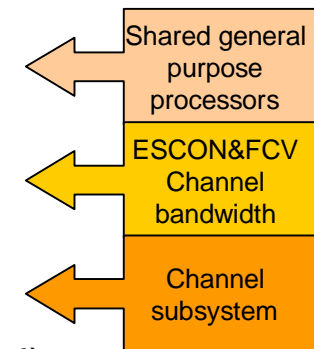
§ This presentation does not include information regarding

- IRD in Basic mode, or in XCFLOCAL mode
 - Refer to References, [2]
- Pricing and software management considerations for workload license charges
 - Refer to References, [3]

Intelligent Resource Director (IRD)

What is IRD?

- § **Set of functions that distribute CPC resources based on business importance**
- § **Problem areas being addressed:**
 - Workloads may change over the course of a day, week, month, year...
 - Business priorities may change over the course of a day, week, month, year...
 - E.g., online vs. batch, production vs. test, workload fluctuations, or periodic work
 - Single static configuration may be sub-optimal to handle different workload mixes
 - Distribute resources based on workload and service level agreements (WLM goals)
 - Reliability problems, e.g. caused by single points of failure
- § **Consists of**
 1. LPAR CPU Management
 - LPAR Weight Management
 - LPAR Vary CPU Management
 2. Dynamic Channel Path Management (DCM)
 3. Channel Subsystem Priority Queuing (CSSPQ)
 - BTW... Not part of IRD: I/O Priority (at control unit level)



LPAR Clusters

§ Scope of IRD management is the LPAR Cluster (LPC)

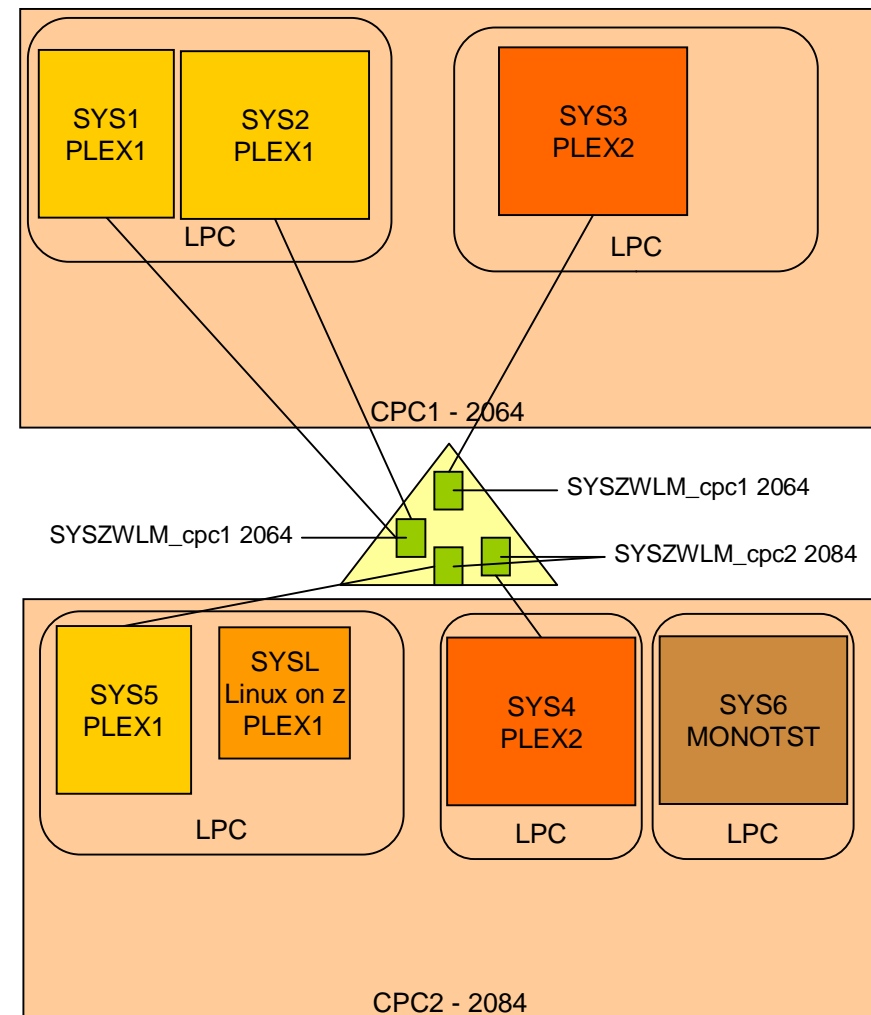
§ LPC:= Set of LPARs on same CPC, which are part of same Sysplex

- For CPU management of zLinux systems, specify the sysplex name as CP management cluster name on *Customize Image Profiles* panel
- For DCM, IOCLUSTER keyword in the IODF must match Sysplex name.

§ **LPARs with dedicated CPs *can* join a cluster**

- But will not be enabled for WLM LPAR Weight and Vary CPU Management.

§ **Multi-image/Sysplex LPCs require a CF structure (except for CSSPQ)**



LPAR Controls at the Support Element

Change LPAR Controls

CPC Name: IP3TVM90
 Last reset profile attempted:
 Input/Output configuration data set (IOCDS): A0 328AF09

Logical Partition	Active	Defined Capacity	Current Weight	WLM Managed	Initial Processing Weight	Minimum Processing Weight	Maximum Processing Weight	Initial Capping	Current Capping	Number of Dedicated Central Processors	Number of Non-dedicated Central Processors	Logical Partition
SCLM1	Yes	0	400	<input type="checkbox"/>	400	0	0	<input type="checkbox"/>	No	0	7	SCLM1
SCLM2	Yes	0	202	<input type="checkbox"/>	202	0	0	<input type="checkbox"/>	No	0	3	SCLM2
SCLM3	Yes	0	202	<input type="checkbox"/>	202	0	0	<input type="checkbox"/>	No	0	3	SCLM3
VM9	Yes	0	0	-	-	-	-	-	No	8	0	VM9
COM1	Yes	0	100	<input type="checkbox"/>	100	0	0	<input type="checkbox"/>	No	0	3	COM1
COM2	Yes	0	80	<input type="checkbox"/>	80	0	0	<input type="checkbox"/>	No	0	5	COM2
COM4	Yes	0	80	<input type="checkbox"/>	80	0	0	<input type="checkbox"/>	No	0	5	COM4
IRL1	No	0	0	<input type="checkbox"/>	1	0	0	<input type="checkbox"/>	No	0	1	IRL1
IRD1	Yes	0	251	<input checked="" type="checkbox"/>	200	50	500	<input type="checkbox"/>	No	0	7	IRD1
IRD2	Yes	0	251	<input checked="" type="checkbox"/>	200	50	500	<input type="checkbox"/>	No	0	4	IRD2
IRD3	Yes	0	102	<input checked="" type="checkbox"/>	102	15	500	<input type="checkbox"/>	No	0	3	IRD3
CFIRD	Yes	0	49	<input type="checkbox"/>	49	0	0	<input checked="" type="checkbox"/>	No	0	1	CFIRD

Processor running time

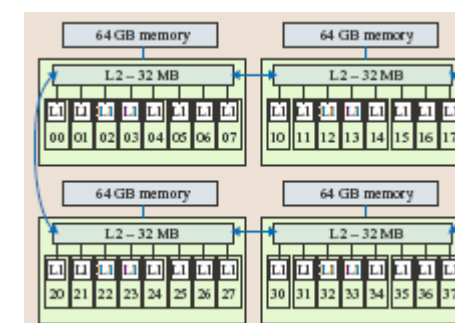
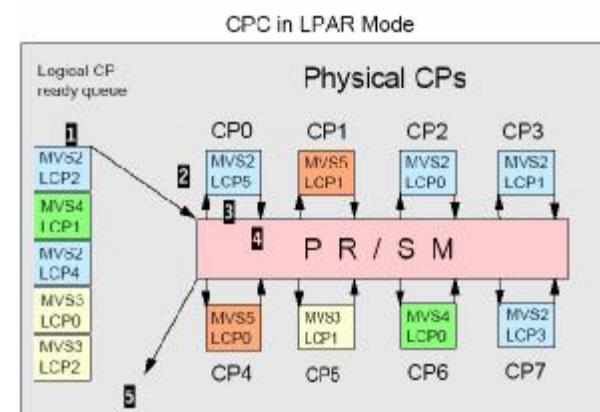
Warning: It is recommended that you select 'Dynamically determined by the system.' Selecting 'Determined by the user' risks suboptimal use of processor resources.

☒ Dynamically determined by the system

Dispatching of Logical Processors

- § z/OS dispatches on logical processors (LCPs)
- § PR/SM Hypervisor dispatches LCPs on physical CPs
- § High number of LCPs may improve responsiveness of an image
- § But large Σ LCPs / PCPs ratio may cause additional overhead
- § Minimum number of LCPs required by an LPAR to consume the capacity that corresponds to its weight share:

$$m_i = \left[\frac{w_i}{\sum_{all LPARs} w_j} \bullet n_{PCPs} \right]$$



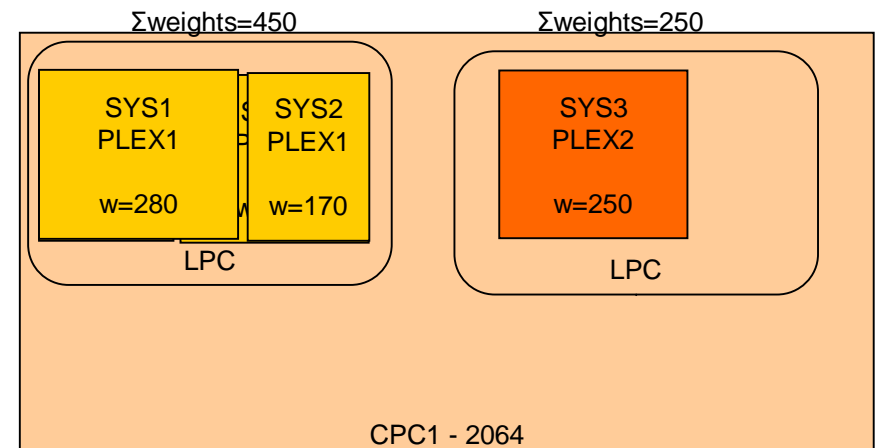
LPAR Weight Management

§ Weights are moved across the LPARs of an LPAR cluster

- The weight of an LPC (sum of “Initial Processing Weight” of all its LPARs) does *not* change.
- Single image LPCs cannot perform LPAR weight management

§ The weight of an LPAR may take any value in the range defined by minimum and maximum processing weight values on the support element.

- LPARs can consume more capacity when CPC utilization allows and no capping is in effect
- Weight will not be increased while LPAR is being capped



Weight Management Controls

§ LPAR Controls

- “WLM Managed” check box
- Initial, Minimum and Maximum processing weights

Change LPAR Controls

CPC Name: IP3TVM90
Last reset profile attempted:
Input/Output configuration data set (IOCD5): A0.328AF09

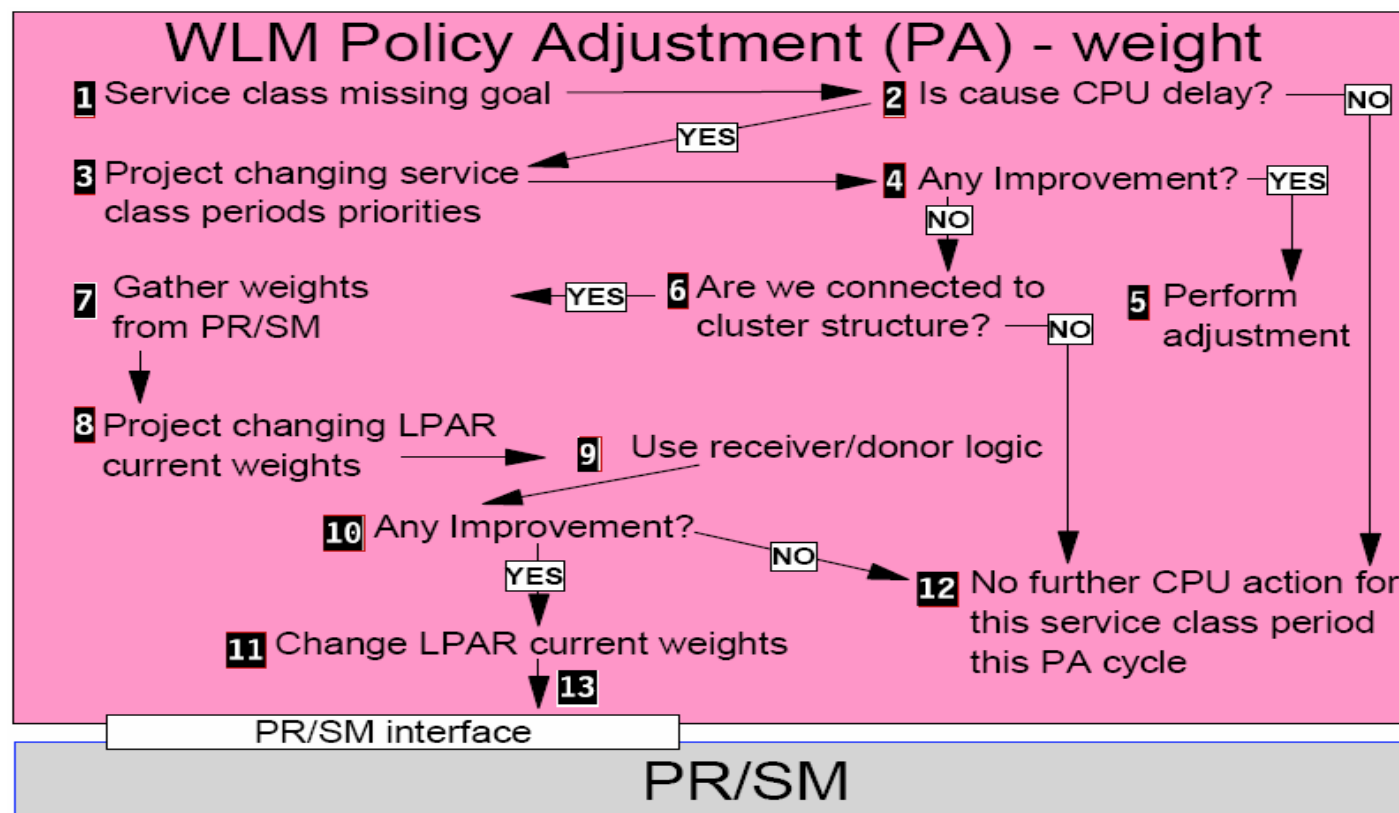
Logical Partition	Active	Defined Capacity	Current Weight	WLM Managed	Initial Processing Weight	Minimum Processing Weight	Maximum Processing Weight	Initial Capping	Current Capping	Number of Dedicated Central Processors	Number of Non-dedicated Central Processors	Logical Partition
SCLM1	Yes	0	400	<input type="checkbox"/>	400	0	0	<input type="checkbox"/>	No	0	7	SCLM1
SCLM2	Yes	0	202	<input type="checkbox"/>	202	0	0	<input type="checkbox"/>	No	0	3	SCLM2
SCLM3	Yes	0	202	<input type="checkbox"/>	202	0	0	<input type="checkbox"/>	No	0	3	SCLM3
VM9	Yes	0	0	-	-	-	-	-	No	8	0	VM9
COM1	Yes	0	100	<input type="checkbox"/>	100	0	0	<input type="checkbox"/>	No	0	3	COM1
COM2	Yes	0	80	<input type="checkbox"/>	80	0	0	<input type="checkbox"/>	No	0	5	COM2
COM4	Yes	0	80	<input type="checkbox"/>	80	0	0	<input type="checkbox"/>	No	0	5	COM4
IRL1	No	0	0	<input type="checkbox"/>	1	0	0	<input type="checkbox"/>	No	0	1	IRL1
IRD1	Yes	0	251	<input checked="" type="checkbox"/>	200	50	500	<input type="checkbox"/>	No	0	7	IRD1
IRD2	Yes	0	251	<input checked="" type="checkbox"/>	200	50	500	<input type="checkbox"/>	No	0	4	IRD2
IRD3	Yes	0	102	<input checked="" type="checkbox"/>	102	16	600	<input type="checkbox"/>	No	0	3	IRD3
CFIRD	Yes	0	49	<input type="checkbox"/>	49	0	0	<input checked="" type="checkbox"/>	No	0	1	CFIRD

Processor running time

Warning: It is recommended that you select 'Dynamically determined by the system.' Selecting 'Determined by the user' risks suboptimal use of processor resources.

Ⓢ Dynamically determined by the system.

Weight Management Algorithm



Weight Management Algorithm

§ LPAR weight management is part of WLM's CPU delay processing:

- Project effect of changing dispatching priorities
- If changing DPs will resolve the CPU delay, implement it
- Else project effect of changing LPAR weights
 - At least 1 min since last weight change?
 - Increase LPAR weight by x%
 - Receiver value?
 - *Select donor LPAR candidate (displaceable capacity at lowest importance)*
 - *Decrease donor LPAR weight by same amount*
 - *Donor impact acceptable?*
 - > ***Implement weight change***

Vary CPU Management

- § Number of logical processors left high as long as free CPC capacity allows for sub-optimal LPAR overhead.
- § When current #LCPs drops out of interval around optimum #LCPs then
 - Config processors offline or online, respectively
 - Interval depends on
 - #current LCPs
 - soft capping state
 - Minimum number of online LPs is determined by maximum of
 - number LPs required to consume current weight,
 - number specified by IEAOPT VARYCPUMIN(*)

(*) Requires z/OS 1.4 or above w/APAR OA09047

```
IEE174I 10.42.49 DISPLAY M 094
PROCESSOR STATUS
ID CPU          SERIAL
00 +           0C9F092064
01 - W
02 - W
03 +           3C9F092064
04 +           4C9F092064
05 +           5C9F092064
06 -
07 -
08 N

CPC ND = 002064.116.IBM.02.000000099F09
CPC SI = 2064.116.IBM.02.000000000099F09
+ ONLINE  - OFFLINE  . DOES NOT EXIST  W WLM-
MANAGED N NOT AVAILABLE
```

Vary CPU Management Controls

§ LPAR Controls

- “WLM Managed” check box
- Initial, Minimum and Maximum processing weights
- Number of Non-dedicated Central Processors

Change LPAR Controls

CPC Name: CPSTVMS0
Last reset profile attempted:
Input/output configuration data set (TOCD5): A0 328AF09

Logical Partition	Active	Default Capacity	Current Weight	WLM Managed	Initial Processing Weight	Minimum Processing Weight	Maximum Processing Weight	Initial Capping	Current Capping	Number of Dedicated Central Processors	Number of Non-dedicated Central Processors	Logical Partition
SCLM1	Yes	E	400	<input checked="" type="checkbox"/>	400	0	0	<input checked="" type="checkbox"/>	No	0	7	SCLM1
SCLM2	Yes	E	200	<input checked="" type="checkbox"/>	200	0	0	<input checked="" type="checkbox"/>	No	0	3	SCLM2
SCLM3	Yes	E	200	<input checked="" type="checkbox"/>	200	0	0	<input checked="" type="checkbox"/>	No	0	3	SCLM3
V319	Yes	E	0	<input type="checkbox"/>	0	0	0	<input type="checkbox"/>	No	8	0	V319
COAM1	Yes	E	100	<input checked="" type="checkbox"/>	100	0	0	<input checked="" type="checkbox"/>	No	0	5	COAM1
COAM2	Yes	E	80	<input checked="" type="checkbox"/>	80	0	0	<input checked="" type="checkbox"/>	No	0	5	COAM2
COAM4	Yes	E	80	<input checked="" type="checkbox"/>	80	0	0	<input checked="" type="checkbox"/>	No	0	5	COAM4
IRL1	No	E	0	<input type="checkbox"/>	0	0	0	<input checked="" type="checkbox"/>	No	0	1	IRL1
IRL4	Yes	E	254	<input checked="" type="checkbox"/>	250	50	250	<input checked="" type="checkbox"/>	No	0	7	IRL4
IRD2	Yes	E	254	<input checked="" type="checkbox"/>	250	50	250	<input checked="" type="checkbox"/>	No	0	4	IRD2
IRD3	Yes	E	102	<input checked="" type="checkbox"/>	100	15	250	<input checked="" type="checkbox"/>	No	0	3	IRD3
CFIRD	Yes	E	49	<input checked="" type="checkbox"/>	40	0	0	<input checked="" type="checkbox"/>	No	0	1	CFIRD

Processor running time

Warning: It is recommended that you select "Dynamically determined by the system." Selecting "Determined by the user" order suboptimal use of processor resources.

0: Technically determined by the system

§ SYS1.PARMLIB

- IEAOPTxx
 - VARYCPU=NO/YES
 - VARYCPUMIN=*number* (*)
- SET OPT=xx

(*) Requires z/OS 1.4 or above w/APAR OA09047

•Operator interaction:

- CPUs that were configured offline by operator will not be varied online by WLM
- Operator cannot config online a processor that is offline by WLM

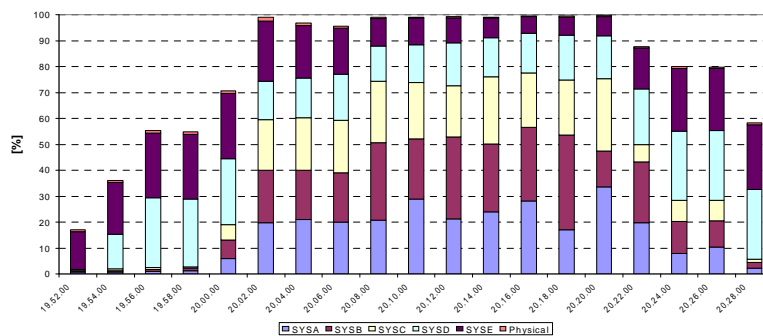
IRD and Shared zAAP/zIIP Processors

- § **zSeries Application Assist Processors (zAAP, also known as IFA processors) are special purpose processors for Java workloads**
- § **They are *not* managed by IRD (same applies for zIIPs):**
 - On z990 and z890, zAAP processors share the same processor pool with ICF and IFL processors
 - But they inherit the weight from the CPs of the LPAR!
See [References](#), [8]
 - IRD weight management does only manage the weights of (general purpose) logical processors – not of zAAPs/zIIPs.
 - IRD Vary CPU management will never config zAAPs/zIIPs on/offline

IRD CPU Management in Action

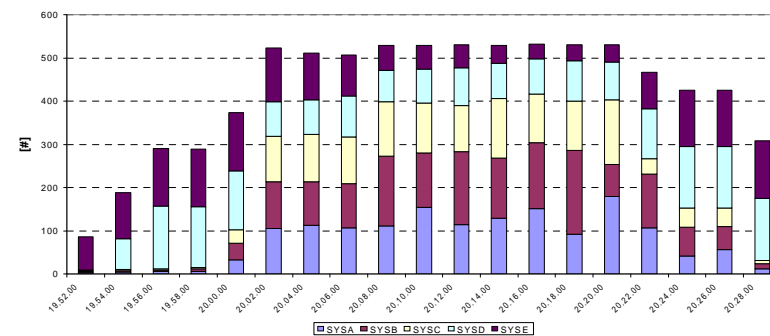
LPAR Utilization
(physical)

Reported Day: 02/21/2003

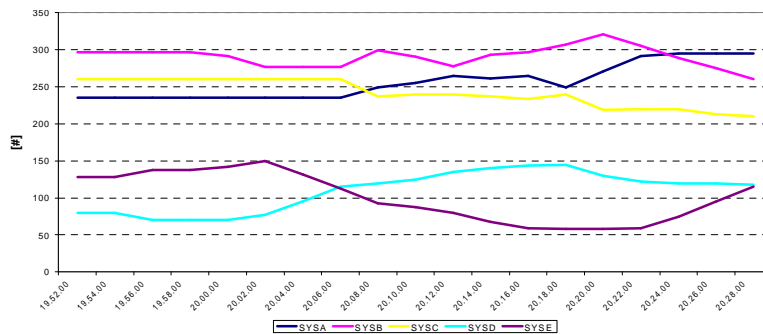


Service Consumed

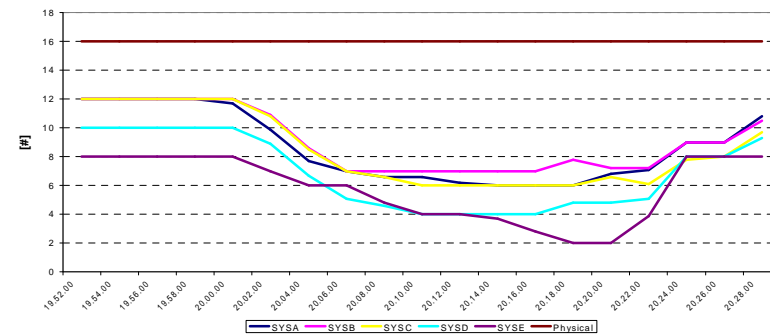
Reported Day: 02/21/2003



LPAR Weights

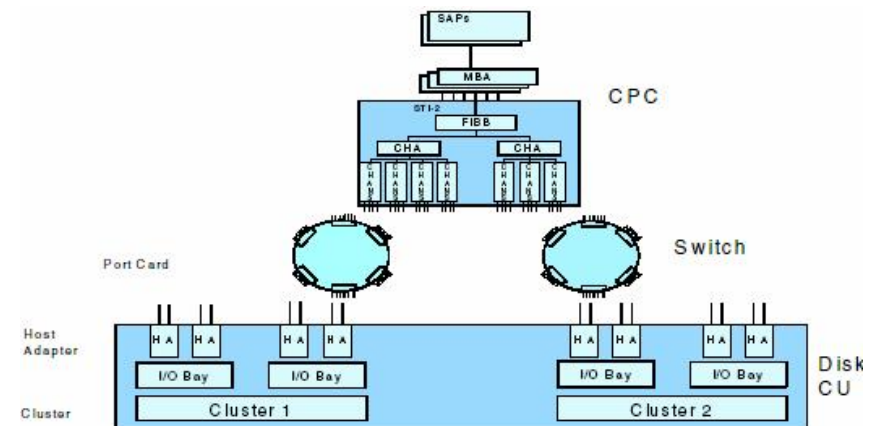


#Logical Processors

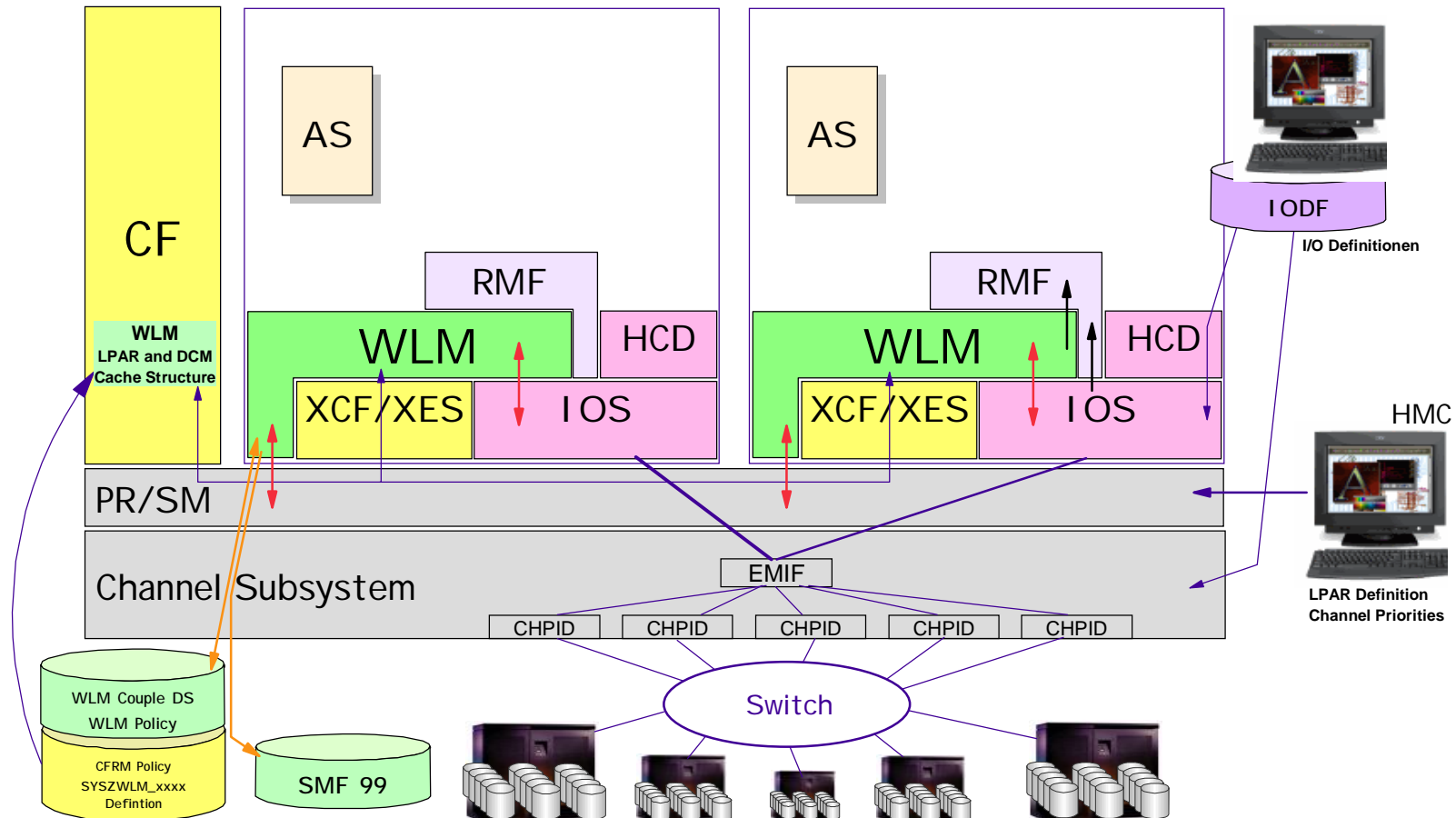


Dynamic Channel Path Management (DCM) Goals

- § **Dynamic Channel Path Management (DCM) lets Workload Manager, in cooperation with IOS, dynamically move channel paths through the ESCON Director from one I/O control unit to another, in response to changes in the workload requirements.**
- § **By defining a number of channel paths as "managed," they become eligible for this dynamic assignment. By moving more bandwidth to the important work that needs it, your DASD I/O resources can be used more efficiently.**
- § **Since DCM considers single points of failures availability in the event of a hardware failure can be improved.**



DCM Componentry



DCM Metrics

§ DCM decisions consider

“I/O velocity”

- Analogously to execution velocity

§ Definition:

$$V_{IO} = \frac{productiveTime}{productiveTime + delays}$$

$$V_{IO} = \frac{connectTime}{connectTime + pendTime - (CUbusy + DEVbusy)}$$

§ Also:

- RAS: Avoid single points of failure
- Complexity: favor “simple” topologies

§ Entire I/O Response time made up of

- Queue time
 - IOSQ
- Pend time
 - Channel busy
 - Switch port busy
 - Device busy from other system
 - CU busy
 - SAP processing
- Connect time
 - Channel data transfer
- Disconnect time
 - Cache miss
 - Reconnect miss
 - Synch. Remote copy

Configuration for DCM

§ Define two static paths to each device

- Minimum is one. DCM will assign a second.

§ Define CHPID as “Managed”

- ESCON or FCV
- Connected to dynamic switch

§ Define control unit with placeholder “*” for CHPID.Link Address

§ Switch device (CUP, port x'FE') needs to be defined

§ Allow for dynamic I/O reconfiguration

§ ...

```
***** Add Channel Path *****
* CBDPCH10
* Specify or revise the following values.
*
* Processor ID . . . : ESERVER      eServer zSeries 900
* Configuration mode : LPAR
*
* Channel path ID . . . . 81 +
* Number of CHPIDs . . . . 7
* Channel path type . . . CNC +
* Operation mode . . . . SHR +
* Managed . . . . . YES (Yes or No) I/O Cluster PLEX1 +
* Description . . . . . Define Managed CHPIDs_____
*
* Specify the following values only if connected to a switch:
*
* Dynamic switch ID . . . 01 + (00 - FF)
* Entry switch ID . . . . 01 +
* Entry port . . . . . 81 +
* F1=Help F2=Split F3=Exit F4=Prompt F5=Reset F9=Swap
* F12=Cancel
*****

***** View Control Unit Definition *****
* CBDCU40 Row 1 of 1 More: > *
* Command ==> _____ Scroll ==> PAGE
*
* Control unit number . : 0400 Control Unit for ESS
* Control unit type . . : 2105 Serial number . . . :
*
* Connected switch.ports : 01.70 01.71 05.90 05.91
*
* ENTER to continue.
*
* Processor Logical -----Channel Path ID . Link Address-----
* ID Address 1---- 2---- 3---- 4---- 5---- 6---- 7---- 8----
*
* ESERVER 95.91 96.90 * * *
* ***** Bottom of data *****
*
*
* F1=Help F2=Split F3=Exit F7=Backward F8=Forward
* F9=Swap F12=Cancel
*****
```

DCM Controls

§ **IOCDs, IODF**

- Managed CHPIDs, CUs, (dynamic) switches

§ **SETIOS DCM=ON/OFF**

- Disables DCM
- Does not revert to initial configuration
- LPAR cluster scope

§ **VARY SWITCH**

- V SWITCH(ssss,pp[-pp]...),DCM=ONLINE|OFFLINE
- LPAR cluster scope

Channel Subsystem I/O Priority Queuing

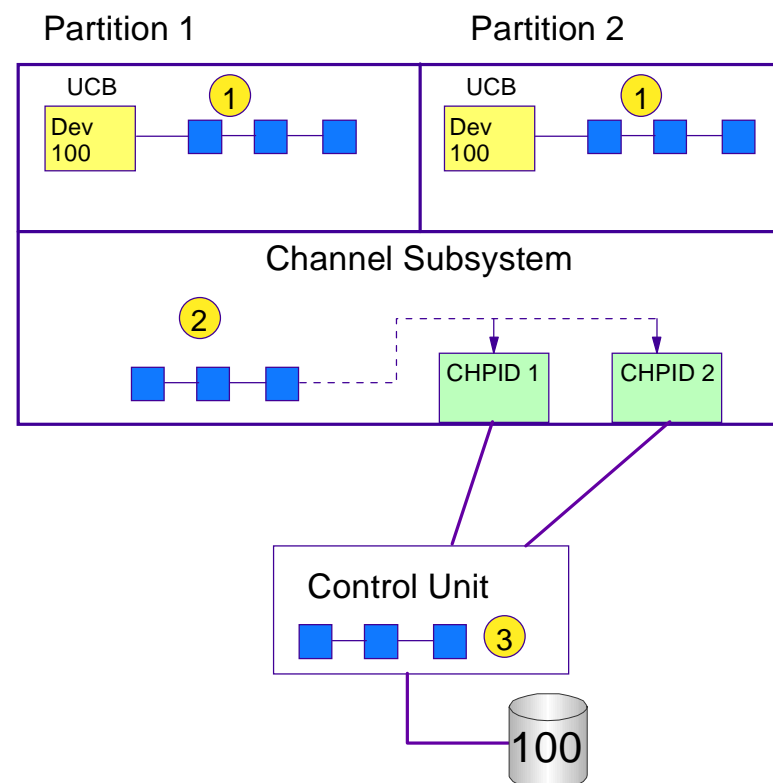
§ **Allows WLM to assign a priority to an I/O request**

§ **Channel subsystem can now use a priority managed queue**

- Prior to this time the channel subsystem used a FIFO queue
- Effective when there is contention in the CSS

§ **Complements Priority Queuing in other parts of the I/O Subsystem**

- IOS UCB Queue
- In CSS waiting for channel
- In control unit



CSSPQ Controls

Change LPAR I/O Priority Queuing

Input/output configuration data set (COCIDS): A0
 Global input/output priority queuing: Disabled
 Maximum global input/output priority queuing value: 15

Logical Partition	Active	Minimum input/output priority	Maximum input/output priority
SCLM1	Yes	0	0
SCLM2	Yes	0	0
SCLM3	Yes	0	0
VM9	Yes	0	0
COM1	Yes	0	0
COM2	Yes	0	0
COM4	Yes	0	0
IRL1	No	0	0
IRD1	Yes	0	15
IRD2	Yes	0	15
IRD3	Yes	0	8
CFIRD	Yes	0	0

§ Image profiles:

- Range of I/O priorities that can be used by the partition
 - WLM maps priority to available range;
 - 0-15 can be specified; range of 8 is sufficient
 - Specify same range for all members of an LPC

§ Reset profile:

- Global switch to activate CSSPQ

Planning for IRD: Prerequisites

Function	Software Prerequisite	Hardware Prerequisite
CPU Management	z/OS R1 and above. Goal mode. CF level ≥ 9 or as required by OS level.	2064 zSeries Processor Driver 36J or later (z800, z990, z890)
CPU Management for non-z/OS LPARs	z/OS R2 and above CF level ≥ 13 or as required by OS level.	Regular CPs (i.e. no IFLs)
DCM	z/OS R1 and above. CF level ≥ 9 or as required by OS level.	ESCON or FICON converted channels. Directors/bridge: 9032-2, -3,-4,-5 w/current LIC IBM 9393 RVA, or IBM 2105 ESS or later.

Agenda

§ Intelligent Resource Director

- CPU Management
- Dynamic Channel Path Management
- Channel Subsystem Priority Queuing

§ WLM Defined Capacity

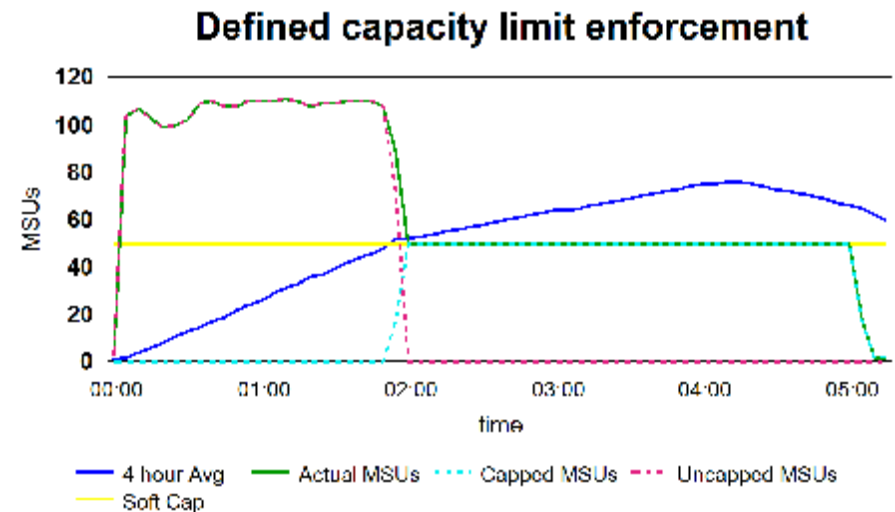
Defined Capacity

- § Sets an upper limit for the sustained capacity that a partition may consume
- § Controlled by WLM , Enforced by PR/SM
 - Within PR/SM precision
- § Specified in “million service units per hour” - MSU
- § Based on *four hour rolling average* computed by WLM
 - WLM computes the service consumed by the partition
 - All intervals before IPL are initialized to 1 MSU/h
- § Capping is only enforced when 4 hour rolling average exceeds the specified limit
 - Cap removed when the average drops below
 - Usage spikes *above* the cap are possible.
 - 4HRA may exceed cap as well!

Change I-PAR Controls

CPC Name: IP3TVM50
Last reset profile attempted:
Input/Output configuration data set (IOCD5): A0 328AF00

Logical Partition	Active	Defined Capacity	Current Weight	WLM Managed	Initial Processing Weight	Minimum Processing Weight	Maximum Processing Weight	Initial Capping	Current Capping	Number of Dedicated Central Processors	Number of Non-Dedicated Central Processors	Logical Partition
SCLM1	Yes	0	400	<input type="checkbox"/>	400	0	0	<input type="checkbox"/>	No	0	1	SCLM1
SCLM2	Yes	0	202	<input type="checkbox"/>	202	0	0	<input type="checkbox"/>	No	0	2	SCLM2
SCLM3	Yes	0	202	<input type="checkbox"/>	202	0	0	<input type="checkbox"/>	No	0	3	SCLM3
VM9	Yes	0	0	-	-	-	-	-	No	8	0	VM9
CDM1	Yes	120	100	<input type="checkbox"/>	100	0	0	<input type="checkbox"/>	No	0	1	CDM1



Capping Mechanisms

§ Capping mechanisms depend on partition weight vs. defined capacity:

Weight vs. defined capacity	How capping is enforced
Weight definition equivalent to defined capacity	PR/SM hypervisor caps at partition weight
Weight equivalent > defined capacity limit	Phantom weight
Weight equivalent < defined capacity limit	Noncontiguous cap pattern

Capping by Phantom Weight

§ Capping at partition weight would be insufficient

- Therefore, introduce a “phantom weight” that is used to pull capacity away from the partition that it is associated with

§ Phantom weight is calculated as

$$\text{PhantomWeight}(i) = \frac{\text{CPC_Capacity}}{\text{DefinedCapacityLimit}(i)} \bullet \text{PartitionWeight}(i) - \sum_{j=1}^{\text{All Active Partitions}} \text{PartitionWeight}(j)$$

- Must be not larger than 1000*number of active partitions
- Consequently there is a lower limit on the defined capacity that can be represented for a given partition weight.

$$\text{DefinedCapacityLimit}_{\text{Min}} \geq \frac{\text{CPC_Capacity} \bullet \text{PartitionWeight}(i)}{1000 \bullet n_{\text{AllActivePartitions}} + \sum_{j=1}^{\text{AllActivePartitions}} \text{PartitionWeight}(j)}$$

- Example:

- Assume 3 partitions with a weight of 500 each on a CPC with 500 MSU.
- Maximum phantom weight is 1000*3=3000
Minimum defined capacity value is 500*500 / (3000+1500)~56

- Then either
 - increase capacity limit, or
 - reduce partition weight

Capping By Noncontiguous Cap Pattern

§ Hypervisor capping at weight equivalent would throttle the LPAR too heavily

§ Therefore,

- Cap the partition for some percentage (x%) of the time,
- Do not limit the partition the remaining time (100-x %),
- such that the average is equivalent to the defined capacity

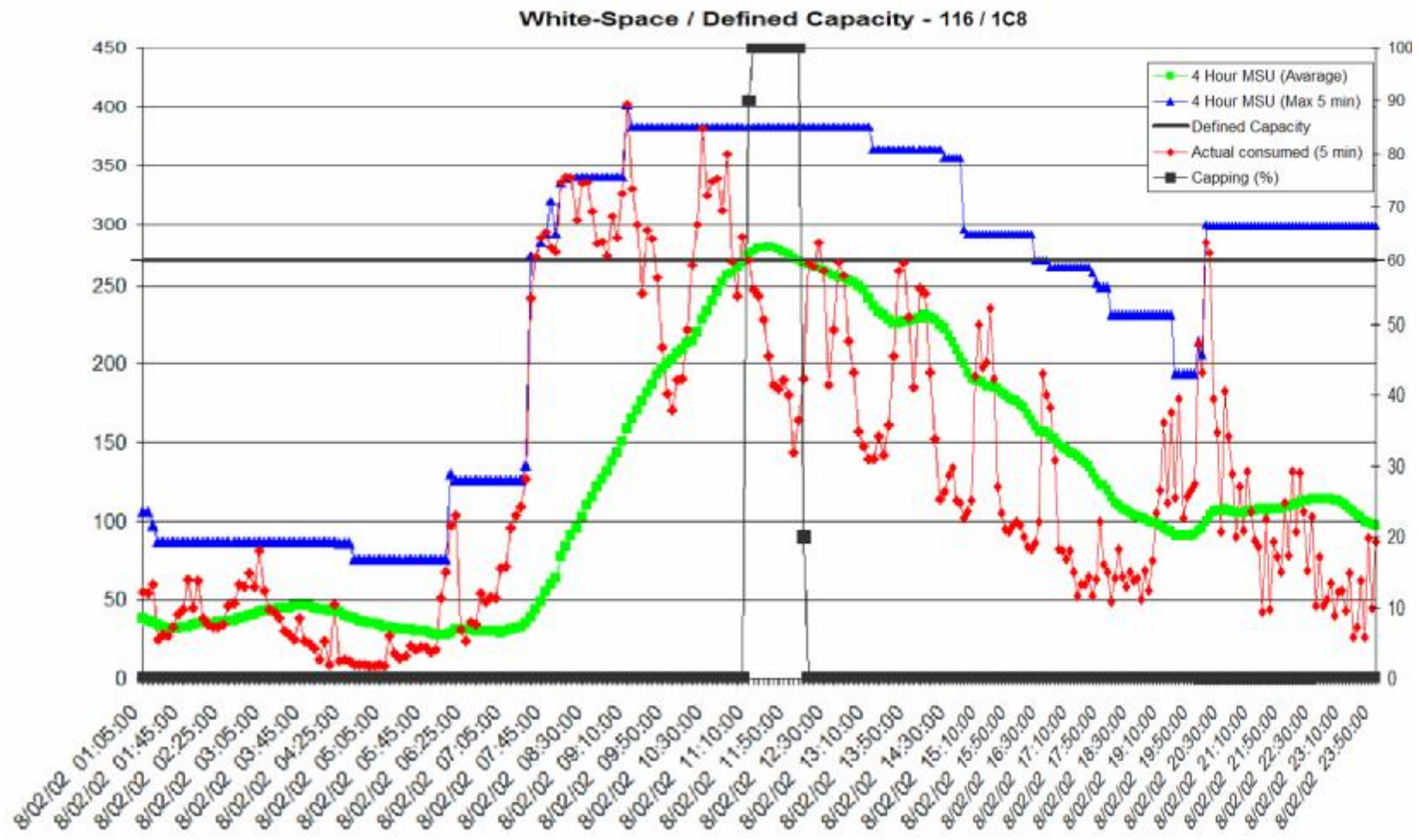
§ Capping will be effective for intervals of 30...600 sec

- If weight is much smaller than defined capacity the partition will be considerably slowed down during those intervals

Defined Capacity and zAAP

- § **Work that executes on zAAPs processors does not contribute to service**
- § **If capping becomes effective HonorPriority=No is enforced**
 - Then zAAP eligible work can run on CPs only below discretionary priority
 - HonorPriority reverts to IEAOPT setting when capping ends

Soft Capping Example



Coexistence of Features

- § **“Initial capping” (*hard capping, LPAR capping*) and “defined capacity” (*soft capping*) are mutually exclusive for any given LPAR.**
- § **“Defined capacity” and IRD functions can be used in most combinations**
 - LPAR weight management should be enabled for Vary CPU management
 - WLM I/O Priority Management must be enabled for CSSPQ

Summary

- § **Advantages and applicability of Intelligent Resource Director**
- § **Technical implementation of Defined Capacity**
- § **Planning required to maximize the benefits**

References

1. **z/OS V1R6.0 MVS Planning Workload Management**
<http://publibz.boulder.ibm.com/epubs/pdf/iea2w150.pdf>
2. **IBM Redbook z/OS Intelligent Resource Director:**
<http://www.redbooks.ibm.com/redbooks/pdfs/sq245952.pdf>
3. **z/OS Planning for Subcapacity Pricing** <http://publibz.boulder.ibm.com/epubs/pdf/e0z2i100.pdf>
4. **zSeries 890 and 990 Processor Resource/Systems Manager Planning Guide**
from IBM Resource Link . E.g. via http://www.ibm.com/servers/eserver/zseries/zos/bkserv/p_sysbooks.html
5. **T. J. Slegel, E. Pfeffer, and J. A. Magee: The IBM eServer z990 microprocessor**
<http://www.research.ibm.com/journal/rd/483/slegel.pdf>
6. **Logical partition mode physical resource management on the IBM eServer z990,**
<http://www.research.ibm.com/journal/rd/483/siegel.html>
7. **z/OS Intelligent Resource Director Planning and Implementation Wizard**
<http://www.ibm.com/servers/eserver/zseries/zos/wizards/ird/irdv1r1/>
8. **z/OS Performance: Capacity Planning Considerations for zAAP Processors**
<http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100417>

Documentation

Documentation



z/OS MVS Planning: Workload Management (SA22-7602)

<http://publibz.boulder.ibm.com/epubs/pdf/iea2w151.pdf>

z/OS MVS Programming: Workload Manager Services (SA22-7619)

<http://publibz.boulder.ibm.com/epubs/pdf/iea2w252.pdf>

Redbook – System Programmer's Guide to: Workload Manager (SG24-6472)

<http://www.redbooks.ibm.com/abstracts/sq246472.html?Open>

Redbook – z/OS Intelligent Resource Director (SG24-5952)

Redbook – Effective zSeries Performance Monitoring Using RMF (SG24-6645)

Internet Links



WLM

<http://www.ibm.com/s390/wlm>

SRM

<http://www.ibm.com/s390/srm>

RMF

<http://www.ibm.com/s390/rmf>

WSC

<http://www.ibm.com/support/techdocs>

Software Pricing

<http://www-1.ibm.com/servers/eserver/zseries/swprice/>

IRD

<http://www-1.ibm.com/servers/eserver/zseries/ird/>

Redbooks

<http://www.redbooks.ibm.com/>

End of Presentation



Thank you very much for your attention