Challenges of MSU Capping w/o Impacting SLAs

St. Louis CMG

Oct 4th 2016 at

Donald Zeunert
BMC Software

Abstract

Numerous IBM capping options to lower MLC costs.

- Difficult to manage and offer SLA risks.
- Free MSU opportunities from some of them.

How can you take advantage of capping and lower the SLA impact risk?

zEnthusiast – Overcome perception mainframe is expensive

Capacity Planners Challenge

Robert De Niro as Simon Silver

Agenda

Capping Objectives

How does 4HRA MLC costs work

- Free MSUs - 4HRA sum vs demand MSUs

Pro/Con of capping options

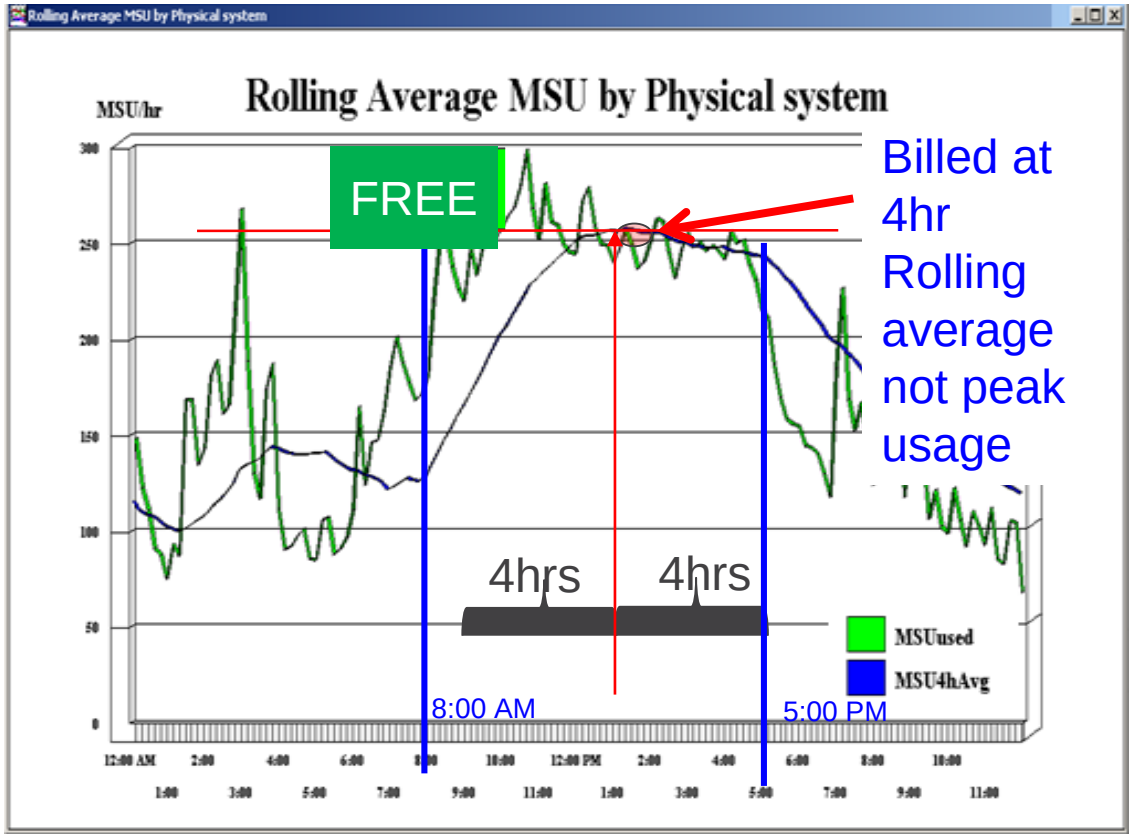
- What are some Risks
- Why did capping fail in the past
  - IBM Improvements to eliminate issues
  - What mechanism are ineffective
  - Which have greatest risks
- How to reduce risks

Capping objectives and options

- Contractual SWLC MSU limit
  - Hard cap, Absolute Cap or Softcap ?
- Lower 4HRA and IBM MLC charges
  - SoftCap
- Control Runaway or unimportant workloads
  - Absolute Cap & Softcap
- Reserve upgrade capacity for later purpose
  - Consistent response, new or growing workload

Note: Hardcap / Initial Cap = Not compatible w/ softcaps

Goal of Usage (4HRA) based MLC



Buy extra capacity - meet SLAs, pay for less than used

"Ideal" workload

- Heavy online with multiple Peaks
- Little or no batch within 4hr of peak
- Batch consumption peak below onlines 4HRA

Q: Is my CPC Ideal?

Typical ISSUEs – 4HRA not when expected

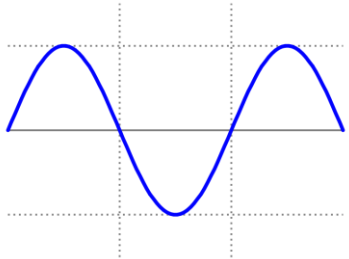
BATCH is source of problem

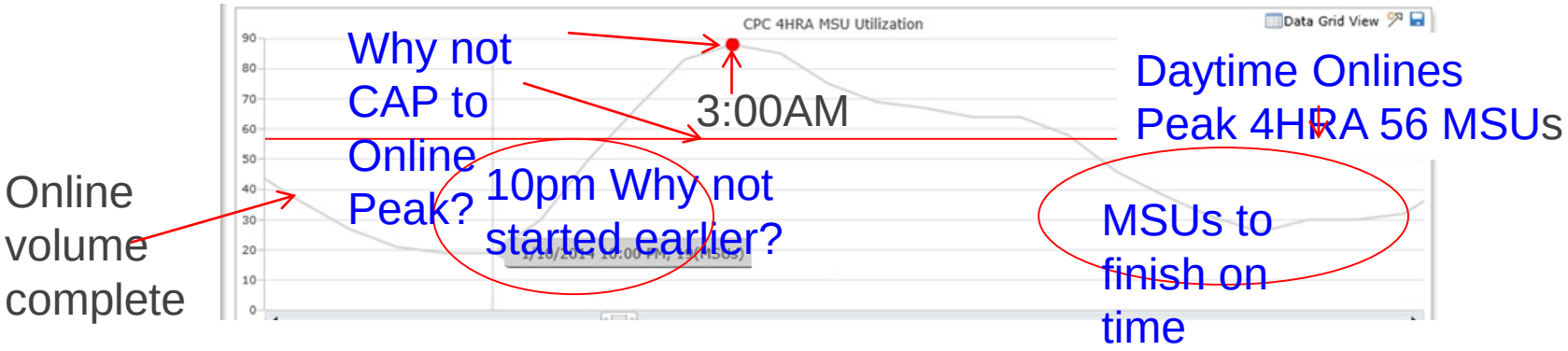
Started too early – averages w/ end of online peak usage

Finishes too late - averages w/ start of online peak usage

Source of peak MSUs is the 4HRA

- Finishes with hours to spare and is not either above condition
  - Needs to be controlled
- Drops then peaks again barely finishes on time





SLA Impact from Capping – WLM

WLM Policy flaws are exposed by resource constraint

- Capping = Resource constraint

WLM Management –

- Does what you told it to do, not necessarily what you want it to do
- Manages to the average of the many (Service class)
- Relative Importance - Do not cause more harm than good

PR/SM LPAR Weights & Hard Cap

Entitlement / SHARE = % of CEC LPAR is guaranteed

$$\text{LPAR Share} = \frac{\text{LPAR Weight}}{\sum \text{Weight of all active LPARs}}$$

- **LPAR share -** NOT a Cap, exceeded if spare MSU
  - Overcommit limited to 100% of the # of LPs
- **Initial / Hard Cap -** limits LPAR to its share
  - Caps even if white space / no 4HRA impact
- **CEC 100% = Sum of entitlements**
  - Only lack of work in uncapped stops from using 100% of CEC

Calculating LPAR Share & MSUs from Weights

LPAR SHARE = LPARs Weight / Sum of Weights

LPAR Name	# LPs	Wgt	Rel Shr	Guar CPs
LPAR1	3	600	60.00%	2.40
LPAR2	2	300	30.00%	1.20
LPAR3	1	100	10.00%	0.40
Total		**1000**		

Processor guarantee = # of Physical GCPs(4) * LPAR Share

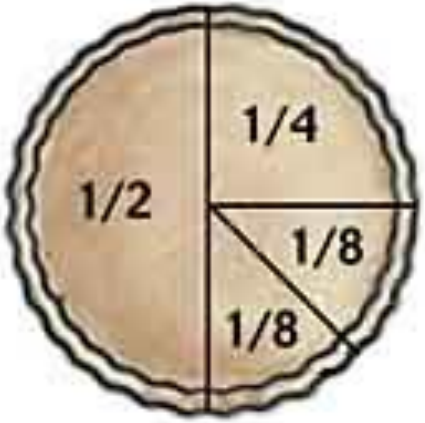
Max MSUs = MSUs per CP * #LPs

Minimum entitled MSUs = MSUs per CP * Guaranteed CP share

Example CEC 2827-504 has 4 GCPs at 70 MSUs each, PR/SM has 6 LPs defined
LPAR 2 Min is 84 MSUs (70 MSUs * 1.20) and
Max is 140 MSUs (70*2) which is 66% more than guaranteed

PR/SM Objectives

CEC at 100%



LPAR Fair Share

If Capping is not controlling all LPARs, capping one LPAR doesn't mean it goes to other one.

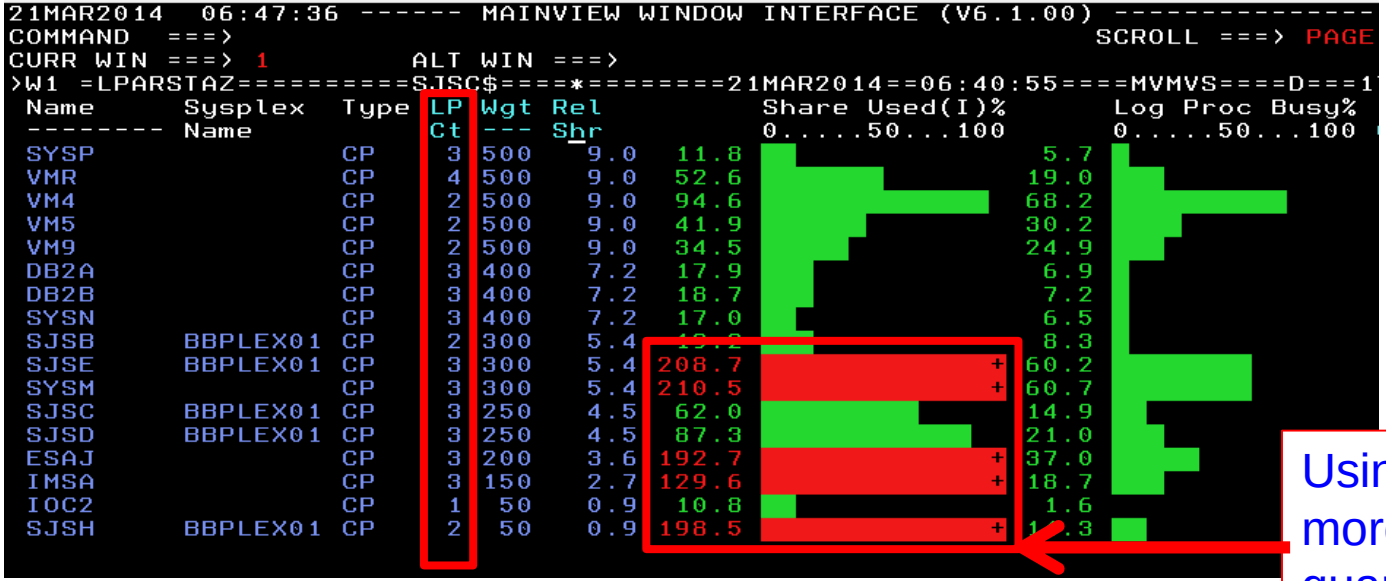
Normal LPAR priorities apply.

Demand MSU Limits

Capping Options	Pros	Cons
PR/SM LPAR Weights	Guarantees minimum share to an LPAR	**DOES NOT LIMIT** Peak or 4HRA MSU consumption
PR/SM LPAR #LPs	Acts like a Hard Cap 100% MSUs of # of LPs. (Granularity is 1 CP)	**DOES NOT LIMIT** Peak or 4HRA MSU consumption, #LPs usually 2-9x minimum / guaranteed share.
Hiperdispatch	Manages # of active LPs improves performance	**DOES NOT LIMIT** Peak or 4HRA MSU consumption
Intelligent Resource Director (IRD)	Manages LPAR Weights to WLM goals and enables Hiperdispatch	**DOES NOT LIMIT** Peak or 4HRA MSU consumption. Can be used in conjunction with Soft Caps
WLM Resource groups	**Hard cap,** controls consumption which may lower 4HRA	**DOES NOT LIMIT** Peak or 4HRA MSU consumption. May prevent low priority work from causing it
	Limits specific WLM Service classes to consume maximum of certain # of SUs in a Sysplex	Doesn't even protect from low importance when Sysplex LPARs on multiple CECs (typical).
PR/SM initial capping (HardCap)	**Hard Cap** - Can be used to control 4HRA costs	**White Space** - wasted when no impact on CEC 4HRA
	Min Guaranteed (CEC MSUs * LPAR Weight %)	Can't use with defined or group capacity
	More granular 1/10 of CP than # LPs	MSU limit increases if other LPAR deactivated
PR/SM Absolute Capping	**Hard Cap** - Can be used to control 4HRA costs. May be used concurrently with defined or group capacity management	**White Space -** wasted when no impact on CEC 4HRA. Needs to be set high & used in conjunction w/ Defined or Group Cap
	Granular 100ths of CP hard limits and independent of LPAR initial weight	Requires EC12 to control how much above 4HRA spikes can be

Logical / Physical – Guaranteed Minimum

Does not create a maximum other than 100% of # of LPs



Even if some LPARs have hard caps, the other LPARs w/o can consume 100% of CEC

Using White Space / more than guaranteed share Impacting 4HRA ?

If no LPARs with work to compete with production batch, can exceed guaranteed share. And therefore easily exceed (2x) daytime 4HRA (3 vs 1.5 CPs)

Everyone's share < 10% of 16 CPs so guaranteed < 1.6 CPs

Why hard Capping failed in the past

MODEL 732 CHANGE REASON=NONE HIPERDISPATCH=NO
H/W MODEL H43

---CPU---		TIME %				LOG PROC
NUM	TYPE	ONLINE	LPAR BUSY	MVS BUSY	PARKED	SHARE %
0	CP	100.00	14.61	54.28	------	18.0
1	CP	100.00	13.00	46.80	------	18.0
2	CP	100.00	10.71	31.82	------	18.0
3	CP	100.00	6.77	18.55	------	18.0
4	CP	100.00	4.22	6.44	------	18.0
5	CP	100.00	4.87	13.16	------	18.0
6	CP	100.00	1.75	2.72	------	18.0
7	CP	100.00	4.54	13.05	------	18.0
A	CP	100.00	4.02	10.40	------	18.0
B	CP	100.00	3.08	6.88	------	18.0
TOTAL/AVERAGE			6.76	20.41		180.0

No HIPERDISPATCH all the CPs got their % CP cut CPU Intensive single TCB workloads suffered heavily.

Not effective use of processor L1 Cache

MODEL 732 CHANGE REASON=NONE HIPERDISPATCH=YES
H/W MODEL H43

---CPU---		TIME %				LOG PROC	
NUM	TYPE	ONLINE	LPAR BUSY	MVS BUSY	PARKED	SHARE %	
0	CP	100.00	89.12	97.67	0.00	100.0	HIGH
1	CP	100.00	87.50	97.83	0.00	80.4	MED
2	CP	100.00	2.51	82.33	96.54	0.0	LOW
3	CP	100.00	1.87	63.68	96.54	0.0	LOW
4	CP	100.00	0.01	-----	100.00	0.0	LOW
5	CP	100.00	0.01	-----	100.00	0.0	LOW
6	CP	100.00	0.01	-----	100.00	0.0	LOW
7	CP	100.00	0.01	-----	100.00	0.0	LOW
A	CP	100.00	0.01	-----	100.00	0.0	LOW
B	CP	100.00	0.01	-----	100.00	0.0	LOW
TOTAL/AVERAGE			18.10	96.92		180.4	

With CPs can run at 100%, CP Intensive workloads not impacted by engine speed

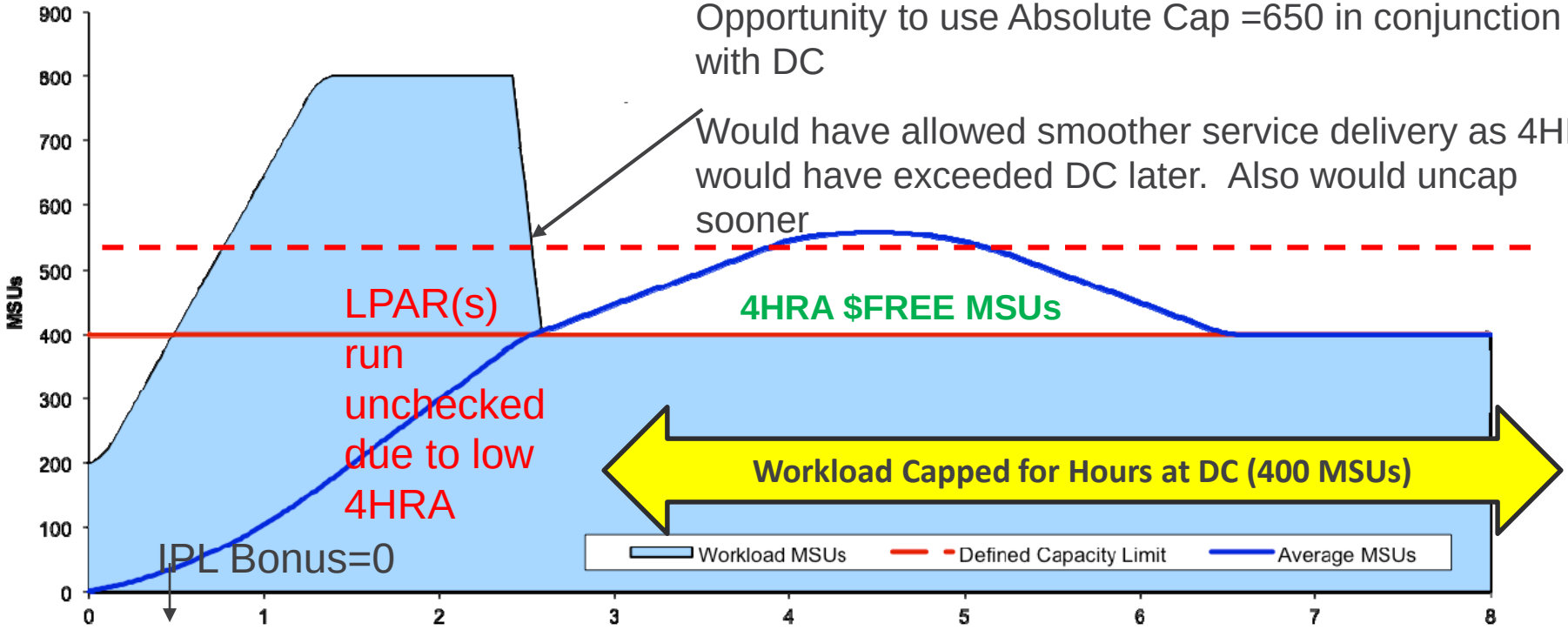
Effective use of processor L1 Cache, may lower CPU secs and (4HRA) billed

4HRA Limits – Soft Capping

PR/SM & WLM Based

Cap Opts	Pros	Cons
Defined Capacity	Allows LPAR to consume MSUs > DC as long as not impact 4HRA. **If exceeds, its not billed**	**White Space** - wasted when no impact on CEC 4HRA. **SLA Risk** - Caps LPARs w/o regard to WLM importance
	Mgd to 4HRA, billing limited. Granular to 1 MSU	**SLA Risk** - Allows MSU Peaks > 4HRA, out perform. Later has to cap to maintain 4HRA.
	Can be used w/ Absolute Capping	**SLA Risk** - Doesn't donate unused capacity
LPAR Group Capacity	Mgd to 4HRA, billing limited. Granular to 1 MSU. Shares capacity among LPARs	**SLA Risk** - Caps LPARs w/o regard to WLM importance, at capacity uses PR/SM shares
	Multiple Capacity groups control max MSUs on subsets of LPARs for VWLC or 3rd party MSUs	**SLA Risk** - Can't share White space that wouldn't impact CEC 4HRA or contracts
	Works w/ Defined Capacity & IRD (as long as group not capped)	**SLA Risk** - Allows MSU Peaks > 4HRA, out perform. Later has to cap to maintain 4HRA.

More FREE MSUs – Exceeding DC or GCL

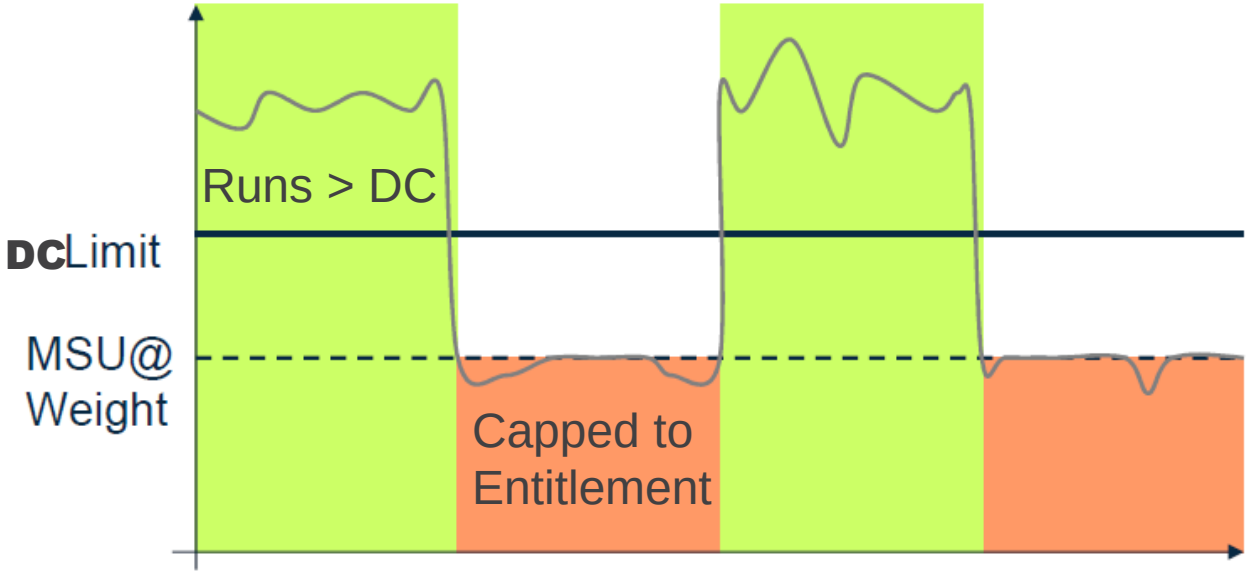


Why Soft Capping failed in the past

Erratic response times from drastic changes to MSUs available

PR/SM Weight (entitlement) vs. Defined Capacity Limit	Hardware / Software level	Selected capping technique
MSU@weight = DCL	Any	Cap at Weight
MSU@weight > DCL	Any	Phantom weight
MSU@weight ≤ DCL	zEC12 GA2 and z/OS V2.1 or later	Negative phantom weight
	Other	Pattern capping

Pattern Capping – DC > LPAR Entitlement



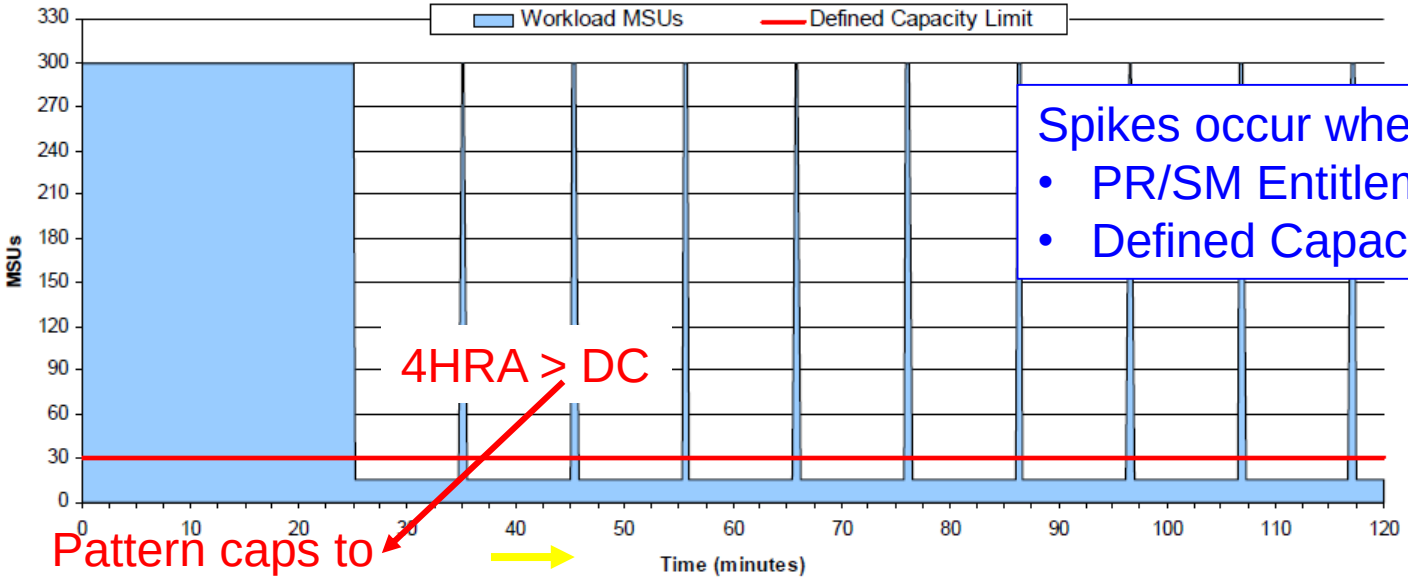
Alternating periods;

- LP capped to MSU@Weight
- LP uncapped

On average the MSU limit is enforced

MSU@weight < Cap Limit

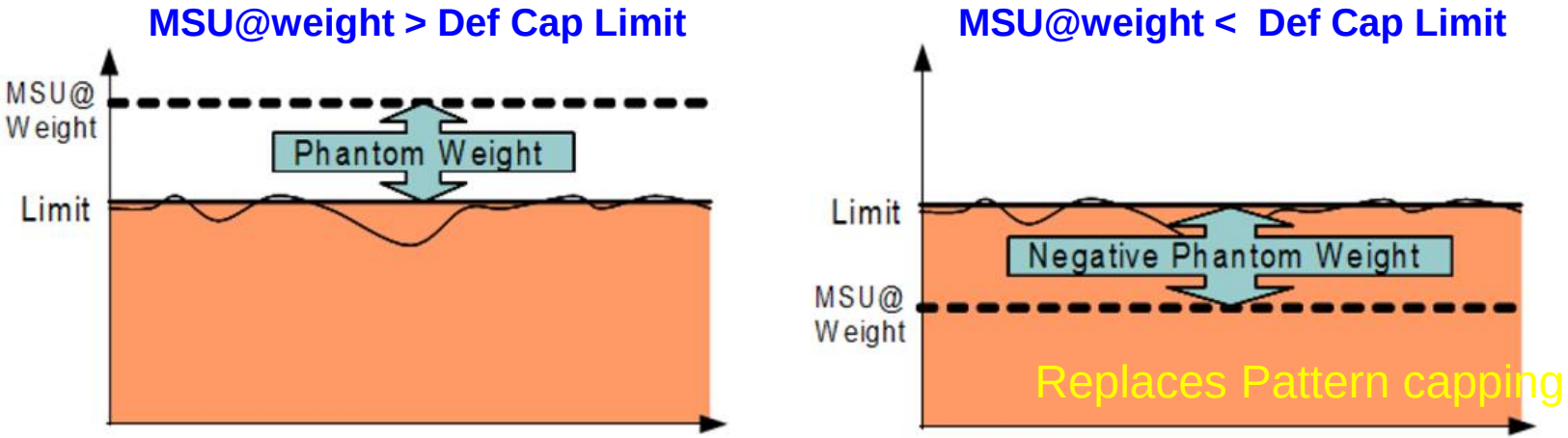
Erratic service delivery – Pattern Capping



Spikes occur when;

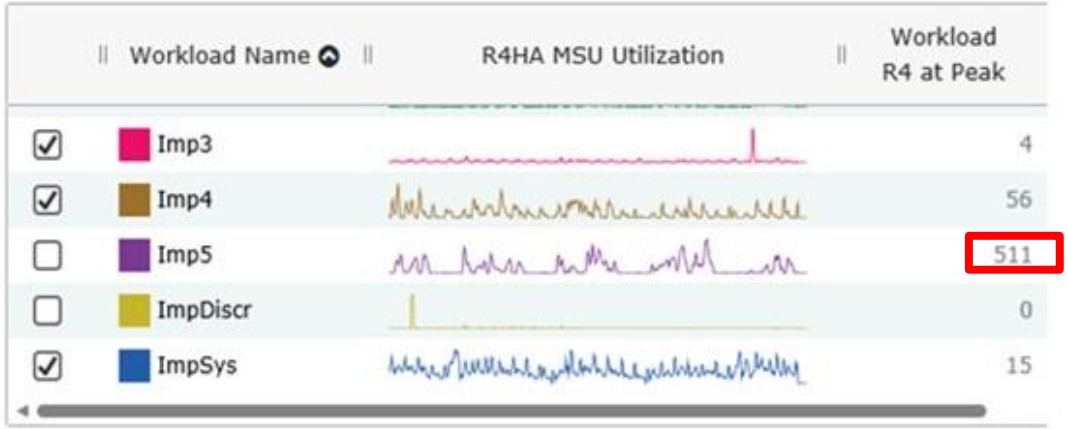
- PR/SM Entitlement is low but
- Defined Capacity limit is high

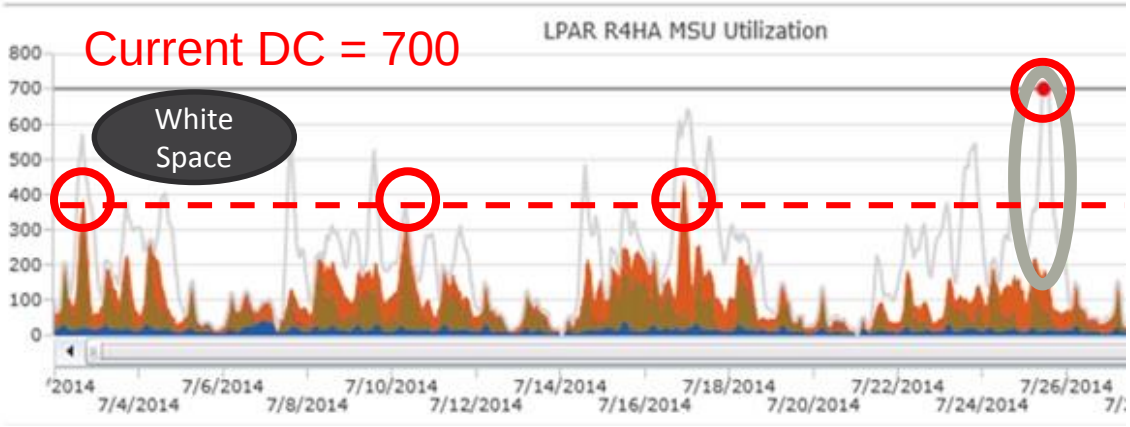
Phantom Weight – Smooth Capping



Negative Phantom Weight – New w/ zEC12 with z/OS V2.1

Static Cap Challenge – Workload variability





White space – constrain w/o benefit

Workload mix at CEC peak not typical

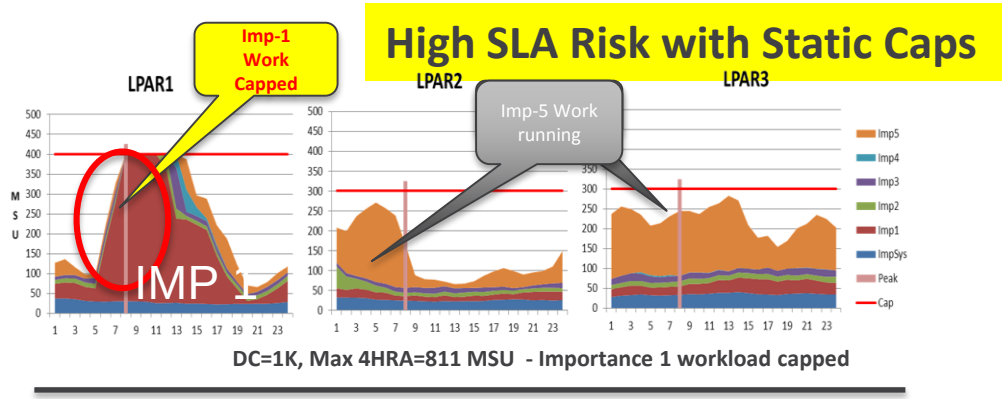
IMP 5 and Discretionary drove peak
Desired DC = 375
But not as much low imp present other times in month.
3 to 4 days cause issue

Benefit of Dynamic vs Static Capping

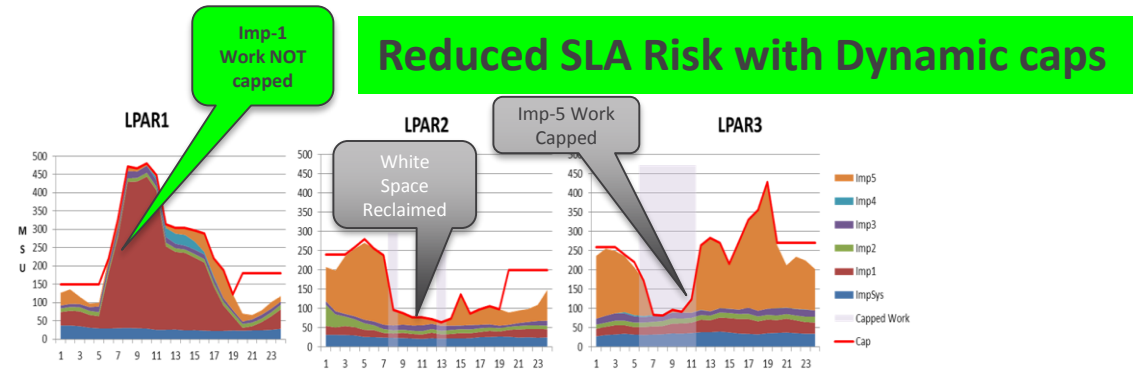
Dynamically Update Capacity

Exploit "white space"

Workload Importance Aware



DC=1K, Max 4HRA=811 MSU - Importance 1 workload capped



MSULIMIT=650 - **Saving of 161 MSU** - No high importance workload capped

Group capping eliminates static DC issue?

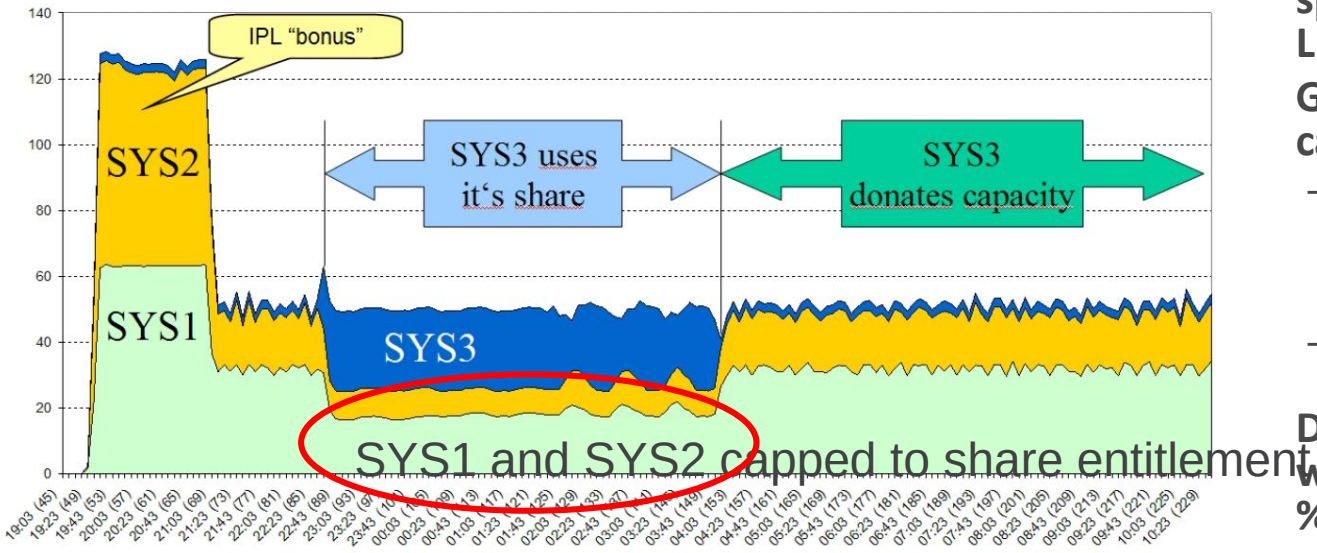
Single GCL - can solve whitespace issues

- Creates SLA and MLC Charges risks (PR/SM Shares)
- IRD - What if Multiple Sysplex / Monoplex on CEC ?

Multiple Capacity groups

- White Space issue same as multiple LPARs w/ Static DCs
- Why were they needed MLC control or SW contract limits?
  - Did you want to steal or give extra?

Capacity Groups – White Space usage



LPARs do not have to be in same Sysplex

Eliminates the White space issue between LPARs

Group members each calculate Group usage

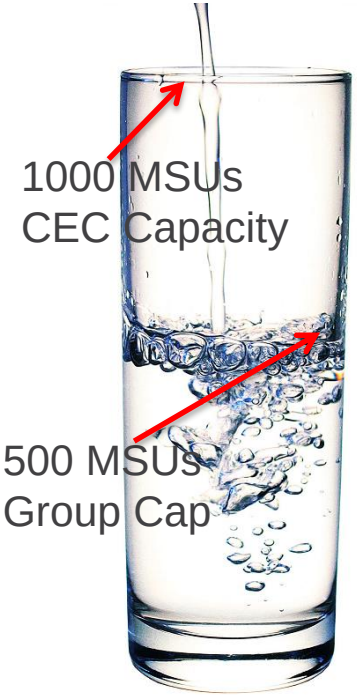
- if exceeds GCL any member detecting caps itself to its entitlement of GCL
- Ignorant to priority of other usage

Donated MSUs shared with receivers based on % GCL share

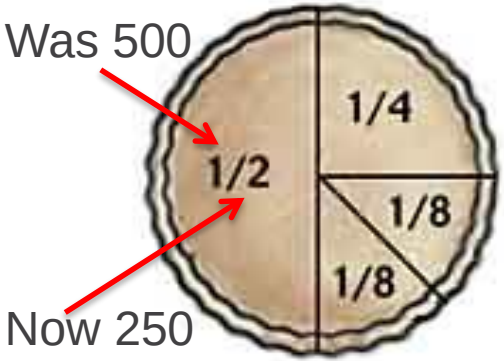
- Up to any LPAR DC

Group Capping – Extreme example

Group MSU Limit



Group PR/SM Entitlement %



Uses PR/SM Weight Ratios

- All LPARs lose equal %
- Need to adjust PR/SM Weights to ensure % of GCL is adequate

Considerations

What was LPARs

- 4HRA Peak?
- Month Peak?
- Original PR/SM Entitled MSUs?

How much low importance was there?

How important is the workload?

Use Intelligent Resource Director (IRD) ?

LPAR cluster - grouping of LPARs on a CEC and in the same Sysplex

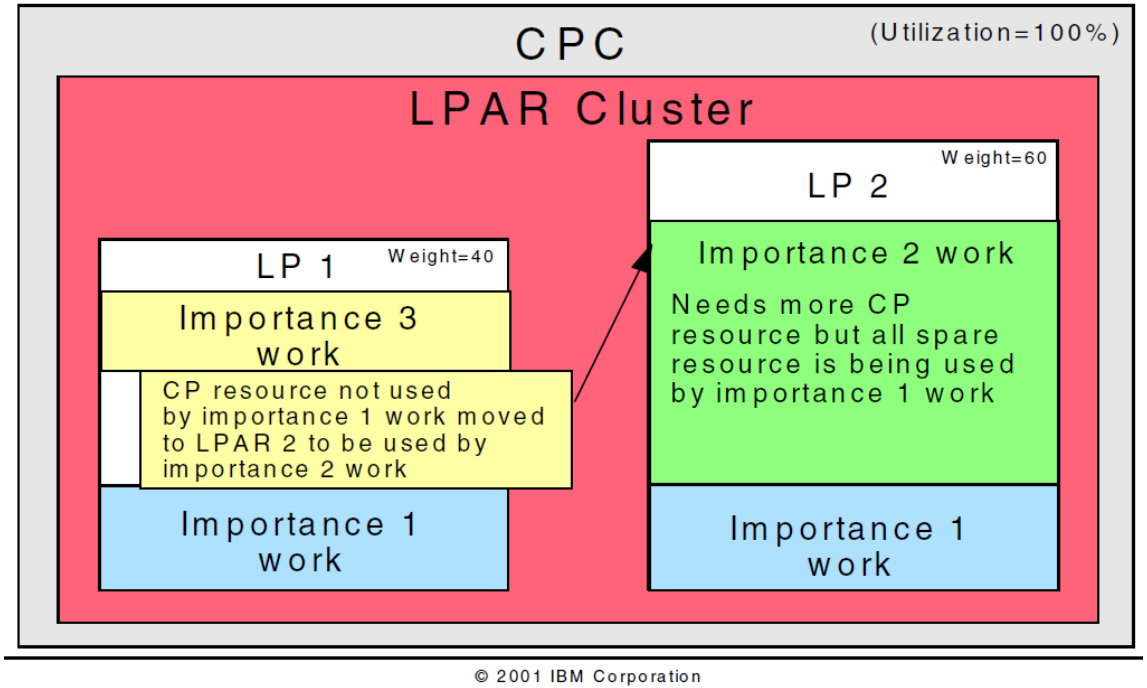
- Multiple LPAR clusters can exist on CEC for multiple Sysplexes
- Monoplexes must be defined as unique Sysplexes
- White Space issues between IRD LPAR clusters

IRD Manages within an LPAR Cluster

- CPU Management for LPARs
  - LPAR weight management within a cluster w/ optional Min/Max per LPAR
  - VARY CPU Management (fixed short CP syndrome, HiperDispatch preferred)
- I/O Performance
  - Dynamic Channel Path Management
  - Channel Subsystem Priority Queuing.

IRD – Reallocation within an LPAR Cluster

Addresses workload routing imbalances within a Sysplex



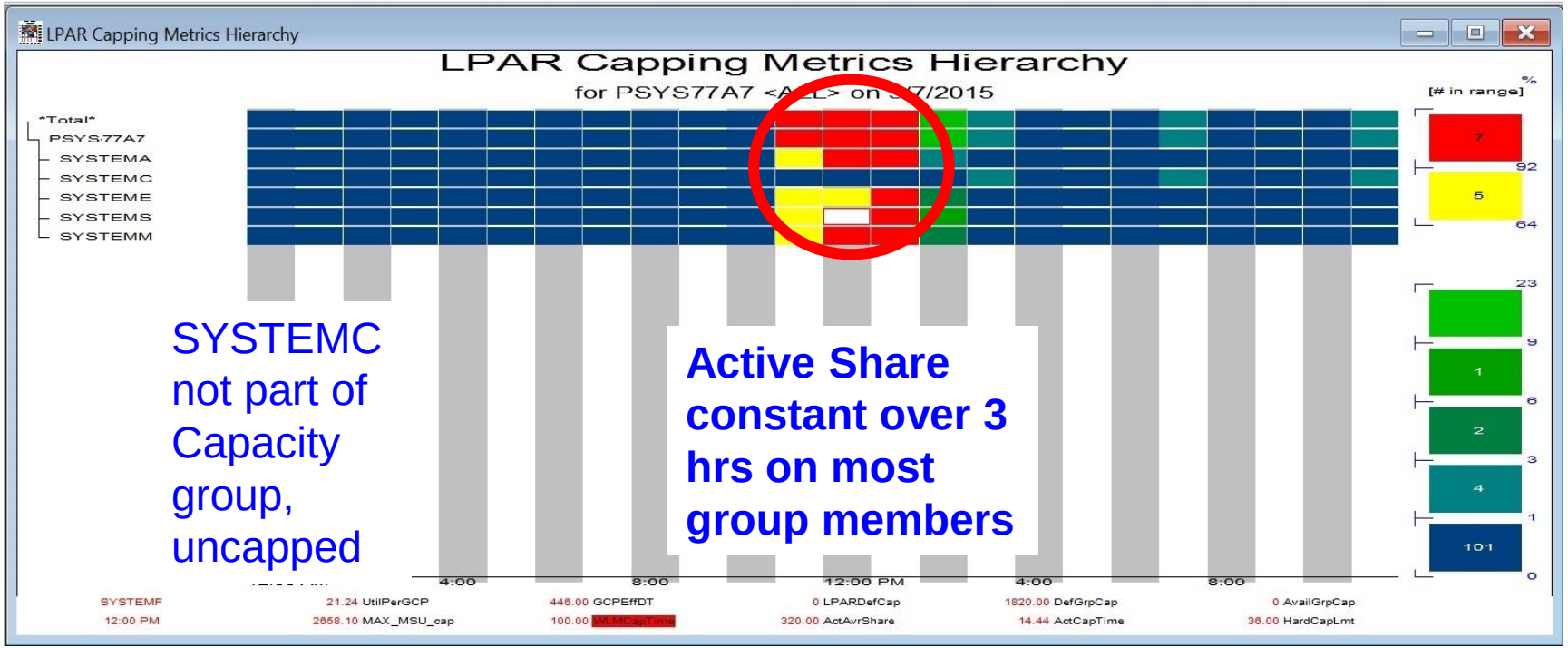
IRD Dynamically adjusts weights steals same % as it reallocates

How many LPARs in 1 Plex do you have on a CPC?

How many Plexes do you have on a CPC?

Redbook - z/OS Intelligent Resource Director SG24-5952-00 (2001)

All Members capped > 92% of the time



Group Cap – w/ IRD Risks –

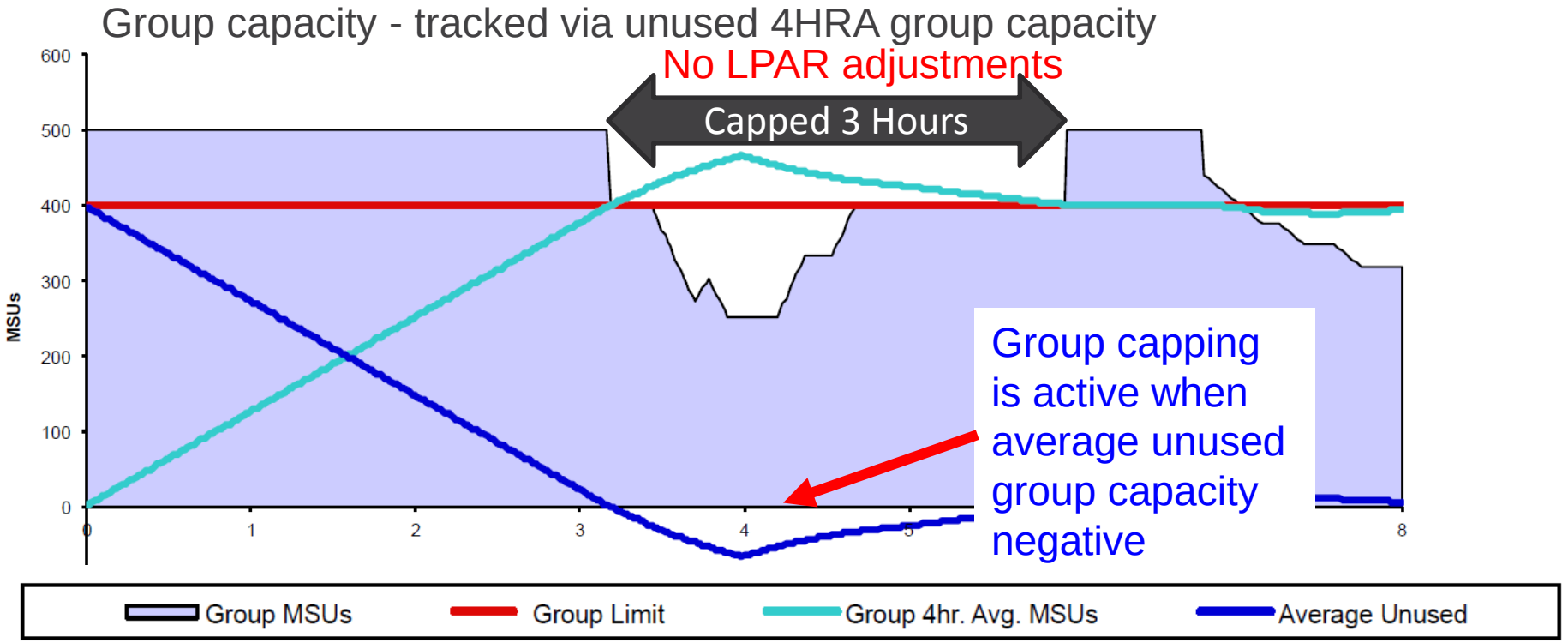
4HRA > GCL – NO IRD Weight Management

- Pre- zOS 2.1 frozen at last setting until 4HRA drops below GCL
- Post zOS 2.1 option to restore to original, frozen until 4HRA drops below GCL

IRD relies on WLM Consistency of LPARs in the LPAR Cluster

- Could be ok as same Sysplex WLM policy

Group Capping -



Group Caps Issues - not necessarily cost cap

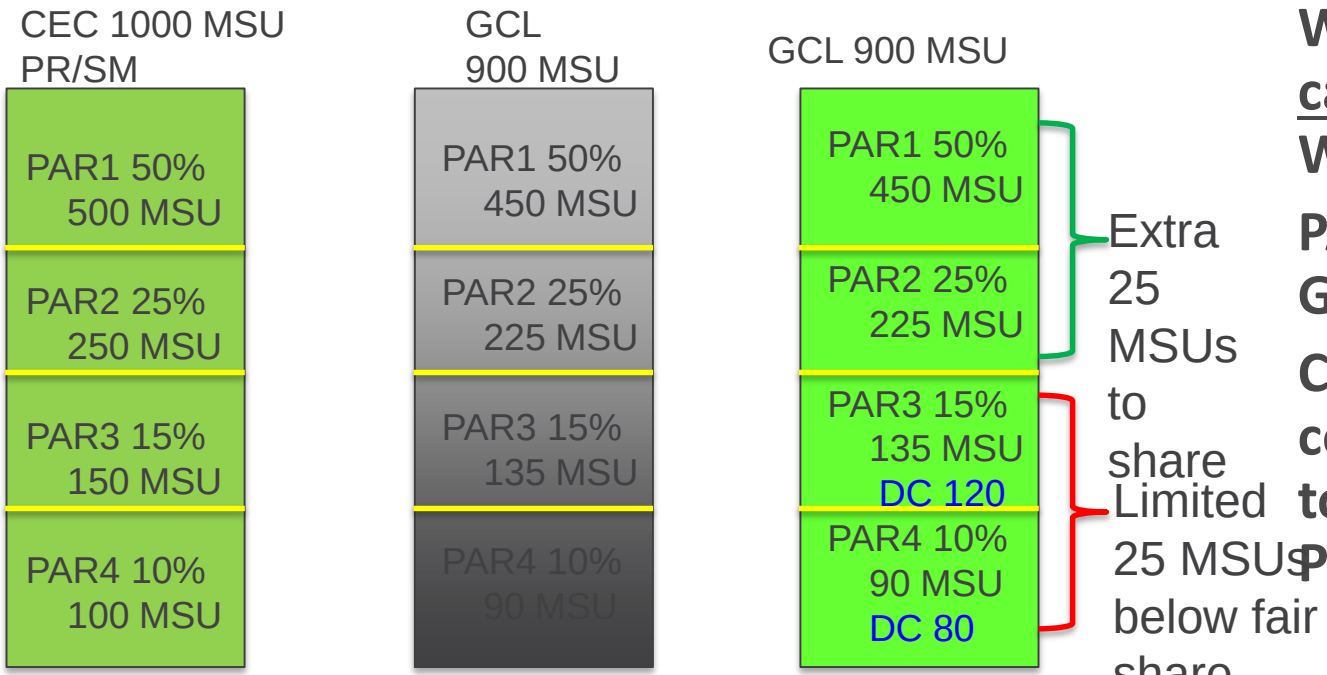
One Large group (typical) – Steal MSUs from cheap (zNALC) and give to expensive (IMS LPAR).

- MSUs capped to max, but cost varies depending on who gets MSUs

Multiple groups – Separate ones for Prod, Test, zNALC

- May provide better cost controls if LPARs in groups have similar $/ delta MSU
- But now back to White Space issues

Using DCs with GCLs



Work on PAR 1/2 <u>can</u> prevent Whitespace

PAR 3/4 donating > GCL reduction %

Can also be used to control max MSU to high $ / MSU Par

Still not SLA aware unless 1 plex

Summary

Capping is safer than before and there is lots of MLC $ to be saved

- IBM hard cap and soft cap have both been improved to provide smoother capping

Static Defined capacities (DCs) are difficult to set low enough

Group capacities (GCLs) have issues, but better than static DCs

Ideally we need mechanism to dynamically adjust DCs / GCLs

- Some customers automate DC changes on schedules
- Several vendors offer solutions based on different criteria to dynamically modify DCs and GCLs

Thank You

Bring IT to Life.™

Contact: Donald_Zeunert@bmc.com