# EPV Performance University 2021

- z/OS Performance Analysis
- WLM Update
- Db2 Performance Analysis

16th June 2021

# Capping and soft capping

# Agenda

- Introduction
- Initial capping
- Absolute capping
- Absolute Group capping
- Soft capping
- Absolute MSU capping

# Introduction

# Introduction

- CPU is the most critical resource in z/OS environments
- It determines most of the HW and SW costs
- Many CPU capping techniques available
- For technical reasons (e.g. to protect important LPARs against other LPARs)
- For financial reasons (e.g. to delay upgrades, to reduce software costs when based on WLC 4-hour rolling average, to limit customers capacity when providing outsourcing services)

# Introduction

| Capping type | Scope | Limit Unit | CP | Controlled by | Enforced by |
|---|---|---|---|---|---|
| Initial (hard) capping | LPAR | %weight | ALL | PR/SM | PR/SM |
| Absolute capping | LPAR | CP fraction | ALL | PR/SM | PR/SM |
| Absolute group capping | Group of LPARs | CP fraction | ALL | PR/SM | PR/SM |
| Defined capacity soft capping | LPAR | MSU 4HRA | GCP | WLM | PR/SM |
| Group capacity soft capping | Group of LPARs | MSU 4HRA | GCP | WLM | PR/SM |
| Absolute MSU soft capping | LPAR | MSU | GCP | WLM | PR/SM |
| Resource group capping | Service class | SU, LPAR share, CP fraction | GCP | WLM | WLM |

All techniques supported in EPV for z/OS

# Initial capping

# Initial Capping

- Initial (hard) capping is the oldest capping technique
- Exploits the PR/SM capping function
- It limits the CPU used by one or more LPAR processor types (independently)
- A capped LPAR running at its cap does not have access to the resources that are not utilized by other LPARs
- Resources that are not used by a capped LPAR can be used by other LPARs

# Initial Capping

- Goals:
  - Limit effect of loops in development LPARs;
  - Limit capacity used during stress tests;
  - Limit capacity (and performance) to what is specified in an outsourcing service contract;
  - …
- Based on LPAR relative weight.
- Expected error is about 1%.
- Control enforced “instantaneously”.

# Initial Capping

- What happens if you cap an LPAR?

| LPAR | Active | Weight | %Weight | Demand | Hcap | %Used |
|---|---|---|---|---|---|---|
| LPAA | YES | 600 | 60% | 50% | NO | 50% |
| LPAB | YES | 300 | 30% | 30% | NO | 30% |
| LPAC | YES | 100 | 10% | 20% | NO | 20% |
| TOTAL | | 1.000 | 100% | 100% | | 100% |

| LPAR | Active | Weight | %Weight | Demand | Hcap | %Used |
|---|---|---|---|---|---|---|
| LPAA | YES | 600 | 60% | 50% | NO | 50% |
| LPAB | YES | 300 | 30% | 30% | NO | 30% |
| LPAC | YES | 100 | 10% | 20% | YES | 10% |
| TOTAL | | 1.000 | 100% | 100% | | 90% |

# Initial Capping

- What happens if you deactivate an LPAR?

| LPAR | Active | Weight | %Weight | Demand | Hcap | %Used |
|---|---|---|---|---|---|---|
| LPAA | YES | 600 | 86% | 50% | NO | 50% |
| LPAB | NO | | | | | |
| LPAC | YES | 100 | 14% | 20% | YES | 14% |
| TOTAL | | 700 | 100% | 70% | | 64% |

- Relative weights change
- Capped LPARs could get more capacity

# Initial Capping

- What about capping the only LPAR active in a CEC ?
- No direct way
- You should create a dummy LPAR

| LPAR | Active | Weight | %Weight | Demand | Hcap | %Used |
|---|---|---|---|---|---|---|
| LPAA | YES | 800 | 80% | 90% | YES | 80% |
| LDUM | YES | 200 | 20% | - | | |
| TOTAL | | 1.000 | 100% | 90% | | 80% |

# Initial Capping

- Initial capping can be used together with HiperDispatch because weights are not changed
- It can’t be used together with IRD weight management because weights are changed
- It can’t be used together with Defined and Group Capacity because they use the same PR/SM capping function

# Absolute capping

# Absolute Capping

- An “absolute capping limit” can be specified:
  - Expressed in terms of 1/100ths of a processor
  - It can be specified per processor type in the image profile and partition controls panel
  - It can be specified independently from the LPAR weight
  - It is insensitive to LPAR (de)activations but sensitive to HW changes
  - Absolute capping may be used concurrently with defined capacity and/or group capacity management; the respective minimum becomes effective

# Absolute Capping

- Initial capping doesn't allow % USED to trespass % TARGET (which is based on LPAR weight)
- % LIMIT is based on the number of LPAR logical processors
- % ABS is based on absolute capping

SYSTEM CEC USAGE PCT SWITCH

CEC USAGE % BY HOUR

| METRIC | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % USED | 35,4 | 26,9 | 30,4 | 22,5 | 22,5 | 18,1 | 9,8 | 14,5 | 19,8 | 36,2 | 37,8 | 34,6 | 30,1 | 26,2 | 33,1 | 33,3 | 35,3 | 32,5 | 33,6 | 36,7 | 34,1 | 35,2 | 28,4 | 25,0 |
| % TARGET | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 54,8 | 48,5 | 54,8 | 48,5 | 48,5 | 48,5 | 54,8 |
| % LIMIT | 64,3 | 64,3 | 64,3 | 64,3 | 64,3 | 64,3 | 64,3 | 64,3 | 64,3 | 50,0 | 50,0 | 50,0 | 50,0 | 50,0 | 50,0 | 50,0 | 50,0 | 64,3 | 64,3 | 64,3 | 64,3 | 64,3 | 64,3 | 64,3 |
| % ABS | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 | 61,6 |

# Absolute group capping

# Absolute Group Capping

- Feature of PR/SM as of z13 GA2, and z13s
- Like LPAR absolute capping but for a group of LPARs
- Specified on the HMC as number of processors with 2 decimal digits (e.g. 3,75)
- Nothing to do with MSU and 4-hour rolling average

# Absolute Group Capping

- SMF metrics for LPAR absolute group capping
  - SMF70HWGr_Name; name of the hardware group this partition belongs to
  - SMF70HWGr_Cap_Limit; if not zero, absolute limit on partition usage of all CPs of the type indicated in SMF70CIX which are members of the same hardware group, in terms of a number specified in hundredths of a CP

# Soft capping

# Soft Capping

- With WLC, major IBM software can be paid based on the highest monthly peak value of the MSU used in a 4-hour rolling average (4HRA)
- WLM continuously measures the LPAR usage collecting data at a 5 minutes interval and calculating the 4HRA (rolling average of 48 intervals)
- RMF collects this info at its interval and writes it in SMF 70 records

# Soft Capping

- The goal of soft capping is reducing the monthly software bill
- Two complementary techniques can be used:
  - defined capacity to limit the MSU usable in the 4-hour rolling average by a single LPAR
  - group capacity to limit the MSU usable in the 4-hour rolling average by a group of LPARs in the same CEC

# Soft Capping

- Defined capacity is a limit set (in HMC) to the capacity, measured by the MSU 4-hour rolling average,  which can be used by an LPAR
- Defined capacity is only applied to CPU (no zIIP)
- CPU usage can spike above the defined capacity limit as long as the MSU used in the 4-hour rolling average remains below it (white space concept)
- When the 4-hour rolling average exceeds the defined capacity limit, WLM will signal PR/SM to cap the LPAR (using initial capping)

# Soft Capping

- WLM is in control but it uses initial capping which can only cap at weight so different tricks have to be used depending on the defined capacity limits compared to weights
- We will try to explain that in the following example

| LPAR | WEIGHT | %WEIGHT | MSU AT WEIGHT | DEF MSU | %DEF |
|---|---|---|---|---|---|
| LPAR A | 300 | 30% | **150** | **100** | 20% |
| LPAR B | 200 | 20% | **100** | **150** | 30% |
| LPAR C | 500 | 50% | **250** | **250** | 50% |
| | 1000 | 100% | 500 | | |

CEC MSU = 500
MSU AT WEIGHT= %WEIGHT *CEC MSU
DEF MSU are the defined capacity limits
%DEF is the percentage MSU allowed by defined capacity limits

# Soft Capping

- LPAR C can be capped without any trick because MSU AT WEIGHT is equal to DEF MSU so initial capping can do its normal job when LPAR C trespasses DEF MSU: capping at weight

| LPAR | WEIGHT | %WEIGHT | MSU AT WEIGHT | DEF MSU | %DEF |
|---|---|---|---|---|---|
| LPAR A | 300 | 30% | **150** | **100** | 20% |
| LPAR B | 200 | 20% | **100** | **150** | 30% |
| LPAR C | 500 | 50% | **250** | **250** | 50% |
| | 1000 | 100% | 500 | | |



# Soft Capping

- In case of LPAR A, MSU AT WEIGHT (150) is greater than DEF MSU (100); initial capping would start at 150 MSUs; so LPAR A would be capped less than needed

| LPAR | WEIGHT | %WEIGHT | MSU AT WEIGHT | DEF MSU | %DEF |
|---|---|---|---|---|---|
| LPAR A | 300 | 30% | **150** | **100** | 20% |
| LPAR B | 200 | 20% | **100** | **150** | 30% |
| LPAR C | 500 | 50% | **250** | **250** | 50% |
| | 1000 | 100% | 500 | | |

# Soft Capping

- To solve the LPAR A issue IBM invented the “phantom weight” (PW) to be assigned to a fictitious LPAR
- PW is calculated in order to make LPAR A MSU AT WEIGHT = DEF MSU

  *PW = (CEC MSU / LPAR A DEF MSU * LPAR A WEIGHT) – SUM(WEIGHT)*

  **PW = (500 MSU / 100 MSU * 300) – 1.000 = 500**
- The fictitious LPAR will get 500 weight so:

  **LPAR A MSU AT WEIGHT = 300 / 1500 * 500 MSU = 100  MSU = DEF MSU**

# Soft Capping

- In case of LPAR B, MSU AT WEIGHT (100) is lower than DEF MSU (150); initial capping would cap LPAR B more than needed when invoked
- To solve the LPAR B issue a “negative phantom weight” is used

| LPAR | WEIGHT | %WEIGHT | MSU AT WEIGHT | DEF MSU | %DEF |
|---|---|---|---|---|---|
| LPAR A | 300 | 30% | **150** | **100** | 20% |
| LPAR B | 200 | 20% | **100** | **150** | 30% |
| LPAR C | 500 | 50% | **250** | **250** | 50% |
| | 1000 | 100% | 500 | | |

# Soft Capping

- Defined capacity limit trespassed in a couple of hours in this example

SYSTEM WLC USAGE

SWITCH

WLC USAGE BY HOUR

| METRIC | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|
| CEC MSU | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 |
| IMAGE MSU | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 |
| GROUP MSU | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 |
| DEF MSU | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 |
| MIN ENT | 36,2 | 36,2 | 36,2 | 36,2 | 36,2 | 36,2 | 36,2 | 36,2 | 36,2 | 36,2 |
| ROLLING 4*HOUR | 26,5 | 33,5 | 40,2 | 41,8 | 43,5 | 45,0 | 46,2 | 50,0 | 51,3 | 51,3 |
| MAX ENT | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 | 51,0 |
| % ACT SOFTCAPP | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 40,6 | 35,8 |
| % WLM SOFTCAPP | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 85,5 | 100,0 |
| GROUP MSU AVA | 444,3 | 336,8 | 227,5 | 171,7 | 153,5 | 128,8 | 133,7 | 113,2 | 105,7 | 109,2 |

# Soft Capping

- Group capacity is an extension of defined capacity and it was designed to limit the MSUs globally used by a group of LPARs in a CEC in the 4HRA
- Group capacity guarantees savings while maintaining the needed flexibility to maximize MSU usage
- A single limit, to the amount of MSUs the Capacity group can use in the 4HRA, has to be assigned in HMC
- It is possible to define multiple capacity groups on a CEC but an LPAR can only belong to one group
- Group capacity and defined capacity can be combined

# Soft Capping

- The following (very old) rules apply when creating a capacity group:
  - LPARs have to be on the same CEC
  - CEC has to be a z9 or a more recent hardware
  - LPARs must run z/OS 1.8 or above
  - LPARs must be assigned shared CPs and wait completion equal to NO
  - LPARs have not be hard capped
- A violation of any of these rules will cause the LPAR to be excluded from the group (4HRA will likely increase)

# Soft Capping

- WLM assigns min and max entitlement values to each LPAR
- If the group has to be capped, the LPARs using more than their min entitlement will be capped

| |
|---|
| Minimum Entitlement; it is the guaranteed MSU share the LPAR can get when in contention; it is calculated as follows:<br>MIN(( WGT x GROUP MSU / SUM(WGT)), DEF MSU ) if DEF MSU GT 0 |
| Maximum Entitlement; it is the maximum MSU share the LPAR can get; it is calculated as follows:<br>MIN( DEF MSU, GROUP MSU ) if DEF MSU GT 0 |

# Soft Capping

- Group capacity limit trespassed in a couple of hours in this example

SYSTEM WLC USAGE

SWITCH

WLC USAGE BY HOUR

| METRIC | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|
| CEC MSU | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 | 2.940,0 |
| IMAGE MSU | 367,0 | 367,0 | 367,0 | 367,0 | 367,0 | 367,0 | 367,0 | 367,0 | 367,0 | 367,0 |
| GROUP MSU | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 |
| DEF MSU | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| MIN ENT | 98,8 | 98,8 | 98,8 | 98,8 | 98,8 | 98,8 | 98,8 | 98,8 | 98,8 | 98,8 |
| ROLLING 4*HOUR | 171,7 | 172,3 | 151,0 | 123,5 | 96,7 | 72,5 | 71,0 | 67,2 | 66,0 | 67,5 |
| MAX ENT | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 | 830,0 |
| % ACT SOFTCAPP | 0,0 | 0,0 | 1,6 | 31,8 | 13,6 | 0,0 | 0,0 | 0,0 | 2,8 | 0,6 |
| % WLM SOFTCAPP | 0,0 | 0,0 | 29,6 | 91,1 | 50,3 | 0,0 | 0,0 | 0,0 | 54,4 | 2,0 |
| GROUP MSU AVA | 298,0 | 155,5 | 19,0 | -28,8 | 2,3 | 23,0 | 32,3 | 26,7 | 0,2 | 15,3 |

# Soft Capping

- It would be interesting evaluating the impact of soft capping on workload performance
- Unfortunately the available measurements are not very useful
- We only have the service/report class CPU delays provided in SMF 72 which include both ‘normal’ CPU delays and soft capping delays

# Soft Capping

- Both soft capping techniques can not be very precise (due essentially to the mobile average mechanism) so it's normal that the MSU used in the 4HRA exceeds the defined and group capacity limits
- Error should be in the 3,6% range (we often found bigger values)
- However user will pay at maximum the capacity limit (exceeding MSUs are a bonus)

# Soft Capping

- Aggressive use of capacity limits could allow to a further reduction in software costs
- However it's important to remember that defined and group capacity limits could be harmful for applications performance (which could be "soft capped" by WLM)
- You normally could accept this risk for low importance workload but not for your "loved ones"

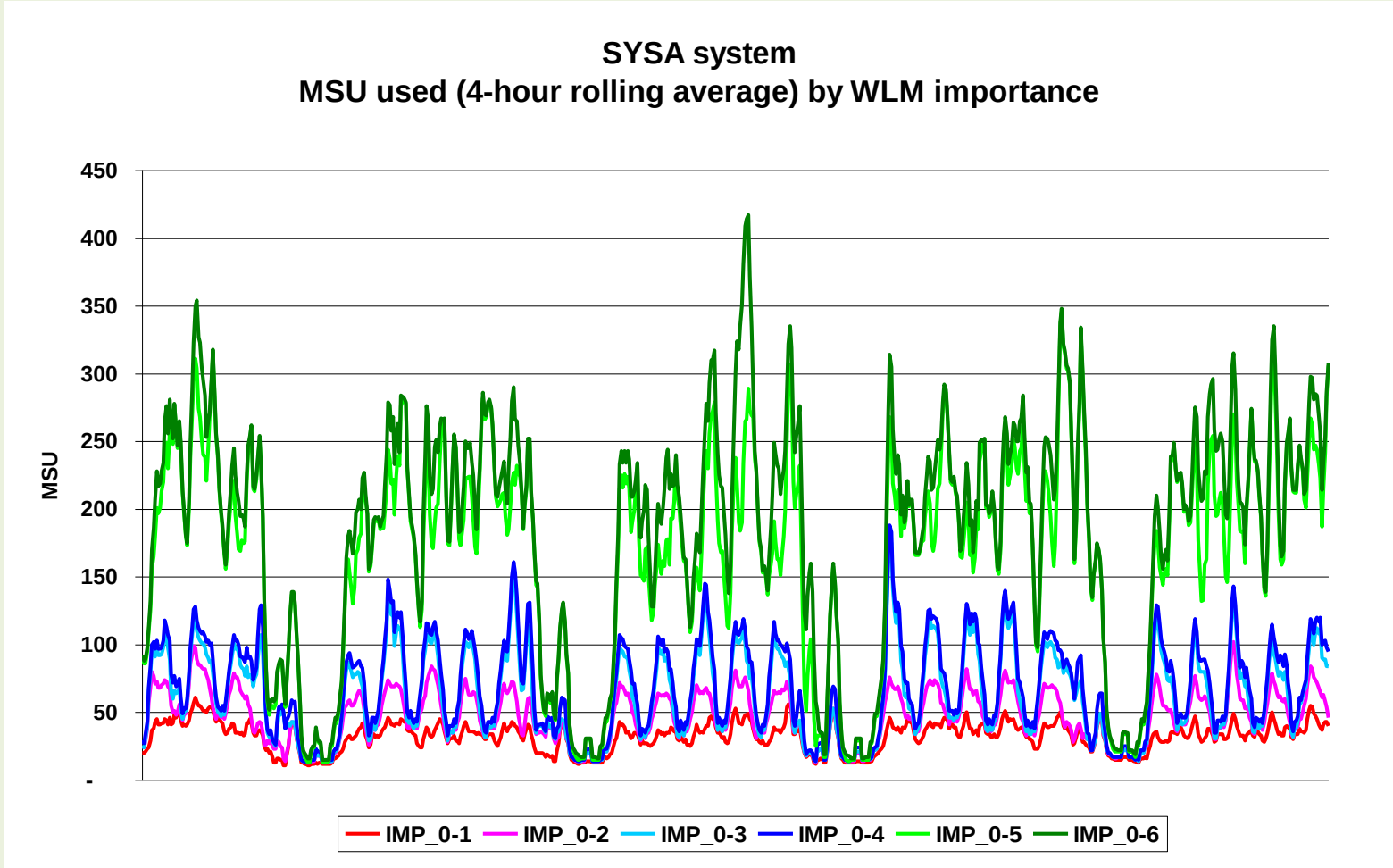
# Soft Capping

- SMF information could be used to:
  - analyze the 4-hour rolling average by WLM workload importance
  - evaluate optimal “defined” and “group” capacity limits
  - estimate possible MSU savings
- In the next slides IMP_0-n accumulates the CPU usage due to SYSTEM, SYSSTC and to all the service class periods having an importance less than or equal to n



# Soft Capping

- IMP_0-1 peaks are slightly above 50 MSUs while IMP_0-4 (which includes everything except importance 5 and discretionary work) never reaches 200 MSUs

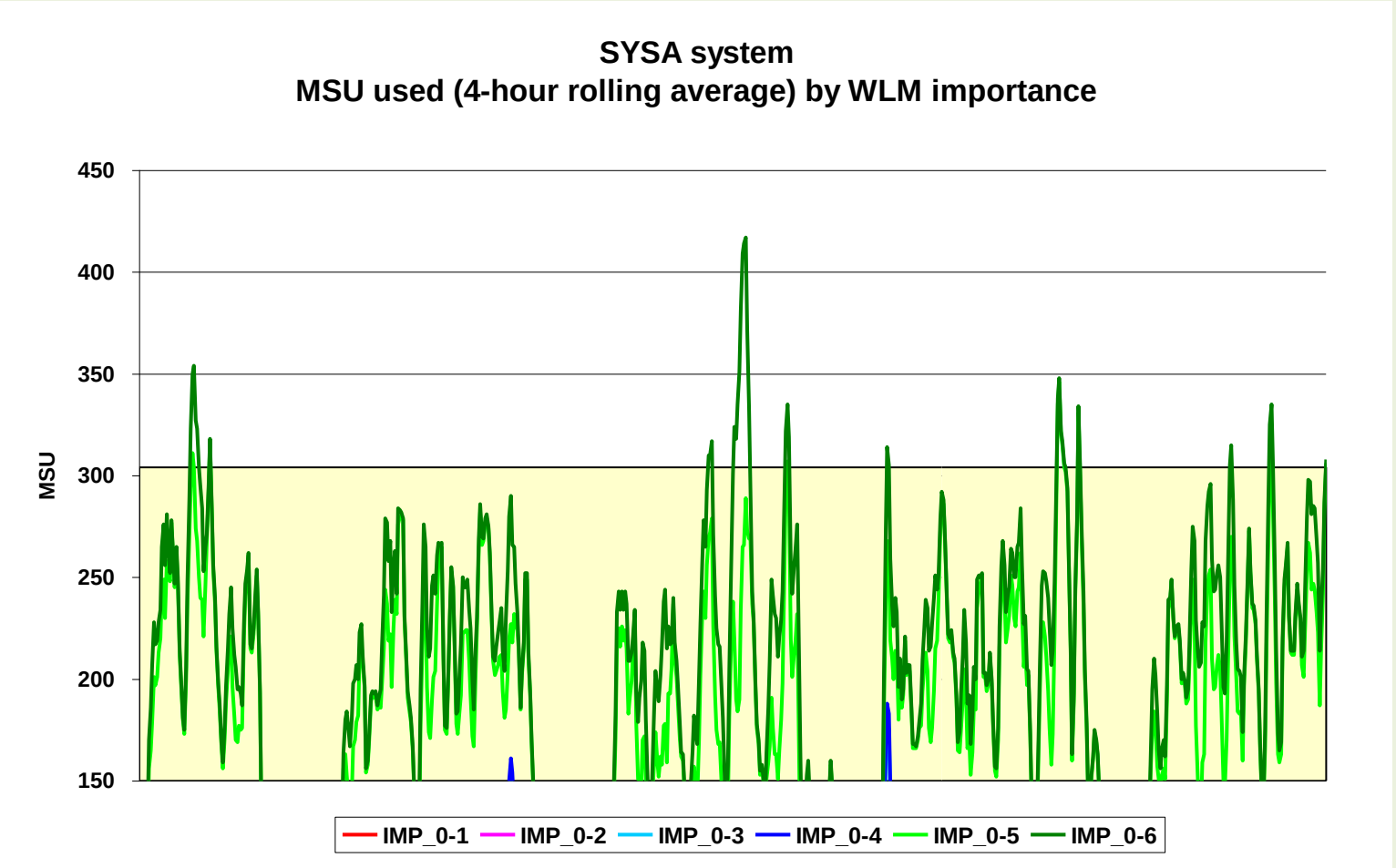


# Soft Capping

- Choosing a percentile as LPAR defined capacity limit you are accepting to degrade some workloads for a certain percentage of the time (5% in case of  a 95 percentile)

| SYSA  4-hour rolling average IMP_0-6 | |
|---|---|
| **PCT** | **MSU** |
| P70: | 236 |
| P75: | 245 |
| P80: | 253 |
| P85: | 265 |
| P90: | 279 |
| **P95:** | **304** |
| P99: | 349 |
| P100: | 417 |

# Soft Capping

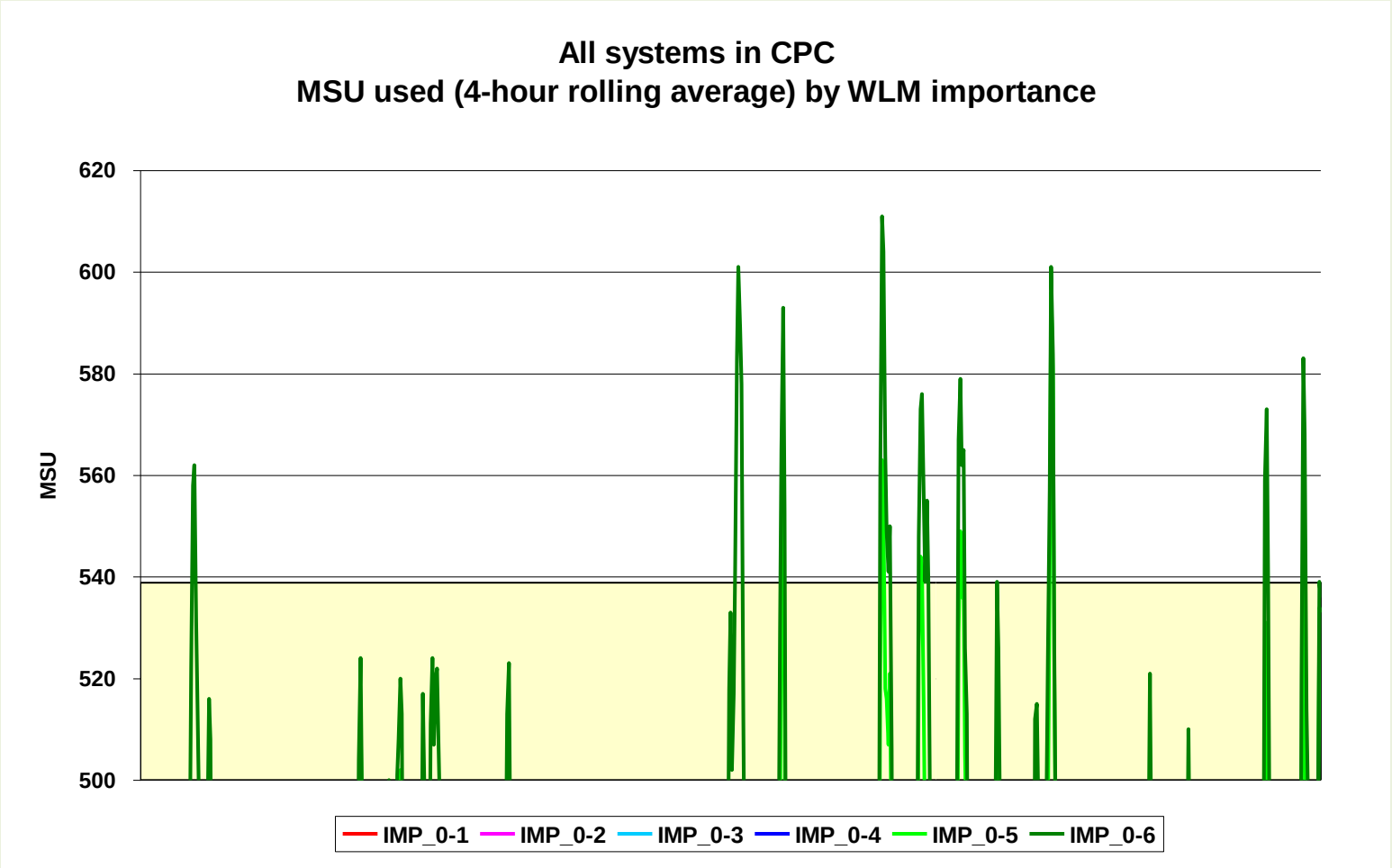


# Soft Capping

- A similar approach has to be used to estimate Group Capacity limit (considering all the LPARs in the CPC)

| CPC 4-hour rolling average IMP_0-6 | |
|---|---|
| **PCT** | **MSU** |
| P70: | 439 |
| P75: | 455 |
| P80: | 471 |
| P85: | 484 |
| P90: | 505 |
| **P95:** | **539** |
| P99: | 583 |
| P100: | 611 |

# Soft Capping



# Absolute MSU capping

# Absolute MSU capping

- Absolute MSU capping:
  - WLM function provided for general purpose processors only
  - ABSMSUCAPPING=YES (NO is default) has to be specified in IEAOPTxx
  - limit is the LPAR defined capacity or LPAR minimum entitlement if group capacity is used
  - independent of 4 hour rolling average consumption; LPAR will always be capped at the limit
  - in the same capacity group can coexist LPARs using AbsMSUcapping=YES and AbsMSUcapping=NO
  - Only SMF metric available is a bit in the SMF70HHF flag

# Questions?