

White Paper

Group Capacity and the Mystery of the Unenforced Limit Fabio Massimo Ottaviani - EPV Technologies



1 Introduction

Most sites pay IBM, and other ISV's, software costs based on the WLC (Workload License Charges) software pricing policy; in this policy, the license fees depend on the CPU usage (measured in MSU's), rather than the machine capacity.

CPU usage is calculated based on a 4-hour rolling average¹; depending on the workload characteristics, this value can be much lower than the power of the machine, which is normally over-sized to guarantee the service levels during a few peak hours.

The "bad news", is that the WLC software license fee is a monthly fee, based on the maximum value of the measured 4-hour rolling average. The complexity of today's systems and workloads, to-gether with human errors, can make it very probable that a company would pay for the full capacity of the machine most of the time.

To guarantee the expected savings, IBM introduced the option to set limits to the MSU which can be used in the 4-hour rolling average:

- by a single LPAR = defined capacity limit
- by a group of LPARs = group capacity limit

The defined capacity limit can be very useful in avoiding certain LPARs, normally running non-business critical workloads from increasing the overall software costs.

1

SEGUS Inc = 14151 Park Meadow Drive = Chantilly, VA 20151= 800.327.9650 = www.segus.com

The sum of the measured 4-hour rolling MSU averages for all the LPARs in the CPC.

The group capacity limit is much more important: it can guarantee that you don't pay more than the limit value, (or more than the sum of the limit values if more than one LPAR group has been created). This is the reason why the majority of the z/OS sites use group capacity limits to protect against the risk of unplanned software costs.

Unfortunately, it can happen that the group capacity limit is not enforced as expected, leading to undesired results.

After a short introduction to Group Capacity concepts, we will discuss this issue based on the experience of one of our customers.

2 Group Capacity Overview

Group capacity limit is an extension of defined capacity, allowing customers to set limits on the MSUs which can be used in the 4-hour rolling average by a group of LPARs².

Users can easily create groups of LPARs, and apply a capacity limit to each of them, by setting the Group Limit and Group Name parameters in the LPAR definitions on the Hardware Management Console.

The following basic rules have to be fulfilled:

- an LPAR can only belong to one group;
- all the LPARs in a group have to run on the same machine.

Additional limitations apply:

- the LPAR must run with shared processors;
- the LPAR must run with wait completion equal "No";
- the operating system must be z/OS V1R8 or higher;
- hardware capping must be used to limit the CPU used by an LPAR.

WLM (Workload Management) uses the definitions of the partitions, and the limits, to calculate a minimum and a maximum entitlement for each LPAR in the group:

Group and defined capacity limits can coexist and work together.

• the minimum entitlement is the guaranteed MSU share the LPAR can get when in contention; it is calculated as:

MIN((WGT X GROUP MSU / SUM(WGT)), DEF MSU) if DEF MSU GT 0;

• the maximum entitlement is the maximum MSU share the LPAR can get; it is calculated as: MIN (DEF MSU, GROUP MSU) if DEF MSU GT 0.

The table in Figure 1 shows an example of group and defined capacity settings as reported in the Group Capacity configurations view³:

	GROUP CAPACITY CONFIGURATION - THU, 25 JAN 2012														
CEC	GROUP	SYSTEM	LPAR- NAME	CEC MSU	GROUP MSU	WEIGHT	DEF MSU	MIN ENT	MAX ENT	САР	OLD Z/ OS	DED	WC=Y		
SER1	Z10ALL	SYS1	LPR1	1329	1010	136	0	137.4	1010	N	N	N	Ν		
SER1	Z10ALL	SYS2	LPR2	1329	1010	717	0	724.2	1010	N	N	N	N		
SER1	Z10ALL	SYS3	LPR3	1329	1010	5	9	5.1	9	Ν	N	N	Ν		
SER1	Z10ALL	SYS4	LPR4	1329	1010	70	126	70.7	126	N	N	N	N		
SER1	Z10ALL	SYS5	LPR5	1329	1010	36	0	36.4	0	N	N	N	N		
SER1	Z10ALL	SYS6	LPR6	1329	1010	36	0	36.4	0	Ν	N	N	Ν		

Figure1

Only one group (Z10ALL) has been created in the SER1 machine. The group capacity limit is set to 1010 MSUs. Defined capacity limits have also been assigned to SYS3 and SYS4 (9 and 126 MSUs) to limit their entitlement.

The four flags at the end of the table indicate that LPAR definitions are compliant to the described group capacity limitations:

- CAP, hardware capping;
- OLD z/OS, z/OS release older than 1.8;
- DED, CPU dedicated;
- WC=Y, wait completion equal "YES".

All the figures present standard view from our EPV for z/OS product.

3 The mystery of the unenforced limit

At a customer site, group capacity is used to control software costs of 6 LPARs running on an IBM 2097-717. Their group and defined capacity definitions are reported in Figure 1.

By looking at the EPV Management Summary view, they realized that something strange had happened in the last month.

MSU USED													
CEC	DATE	INST	USED	BASELINE									
SER1	2012-01	1329	1070	1010									
SER1	2011-12	1329	933	1010									
SER1	2011-11	1329	973	1010									
SER1	2011-10	1329	965	985									
SER1	2011-09	1329	913	085									
SER1	2011-08	1329	904	970									
SER1	2011-07	1329	911	970									
SER1	2011-06	1329	956	970									
SER1	2011-05	1329	920	950									
SER1	2011-04	1329	940	950									
SER1	2011-03	1329	883	950									
SER1	2011-02	1329	952	950									
SER1	2011-01	1329	944	950									

Figure 2

The monthly peak of the MSU, used in the 4-hour rolling average, (USED), in January 2012, is 60 MSUs more than the group capacity limit (BASELINE). The soft capping algorithms used by defined and group capacity can't be extremely precise, so it may happen that the MSUs used are slightly more than the limits, (see also February 2011 in the above figure).

This is an advantage for the customer, who doesn't have to pay for these extra MSUs; they will be charged taking into account the minimum value of the limit set and the MSU used.

However, 60 MSUs seemed a bit high to be considered normal soft capping "imprecision". So they decided to deepen their investigation.

	CEC: SER1 BY GROUP														
					Z10	DALL	NOLIMIT								
DATE	TYPE	MODEL	MSU	TOTAL	MSU LIMIT	MSU USED	MSU USED								
2012-01	2097	717	1329	1070	1010	975	95								
2011-12	2097	717	1329	975	1010	933									
2011-11	2097	717	1329	913	1010	973									
2011-10	2097	717	1329	873	985	965									
2011-09	2097	717	1329	865	985	913									
2011-08	2097	717	1329	913	970	904									
2011-07	2097	717	1329	867	970	911									
2011-06	2097	717	1329	861	970	956									
2011-05	2097	717	1329	856	950	920									
2011-04	2097	717	1329	879	950	940									
2011-03	2097	717	1329	728	950	883									
2011-02	2097	717	1329	823	950	952									
2011-03	2097	717	1329	883	950	944									

Figure 3

An additional NOLIMIT group, which used 95 MSUs, is reported in the WLC by Group view, (see Figure 3), besides the Z10ALL group, but only in January 2012.

Drilling down to the day level, the problem seems to be restricted to January 26th, which is also the peak of the month.

MSU 1329 1329 1329	TOTAL 704 712	Z10ALL 704 712	NOLIMIT
1329 1329 1329	704 712	704	
1329 1329 1320	712	712	
1329	745	/ 12	
1720	/45	745	
1529	419	419	
1329	823	823	
1329	929	929	
1329	1070	975	95
1329	964	964	
1329	816	816	
1329	767	767	
1329	350	350	
1329	784	784	
1329	907	907	
1329	882	882	
1329	943	943	
1329	867	867	
1329	786	786	
1329	336	336	
1329	630	630	
1329	841	841	
1329	761	761	
1329	787	787	
1329	851	851	
1329	761	761	
1329	318	318	
1329	661	661	
1329	740	740	
1329	771	771	
1329	816	816	
1329	785	785	
1329	792	792	
	1329 1329 <td>13297451329419132982313299291329107013299641329767132976713297841329907132988213299431329867132978613297861329786132978613297871329761</td> <td>13297457451329419419132982382313299299291329107097513299649641329816816132976776713297847841329907907132988288213299439431329867867132978678613297867861329786786132978181113297617611329761761132976176113297617611329761761132976176113297617611329761761132976176113297617611329761761132976176113297617611329761761132976176113297407401329740740132974074013297717711329785785132978578513297857851329785785</td>	13297451329419132982313299291329107013299641329767132976713297841329907132988213299431329867132978613297861329786132978613297871329761	13297457451329419419132982382313299299291329107097513299649641329816816132976776713297847841329907907132988288213299439431329867867132978678613297867861329786786132978181113297617611329761761132976176113297617611329761761132976176113297617611329761761132976176113297617611329761761132976176113297617611329761761132976176113297407401329740740132974074013297717711329785785132978578513297857851329785785

Figure 4

	CEC : SER1 - WORKLOAD: Z/OS - 4 HOUR MOVING AVG BY HOUR - THU, 26 JAN 2012																											
GROUP	SYSTEM	TYPE	MODEL	MSU	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Z10ALL	SYS1	2097	717	1329	69	74	71	51	34	31	37	48	59	69	71	67	63	57	57	55	56	69	58	58	49	41	43	55
Z10ALL	SYS2	2097	717	1329	708	664	648	663	660	646	648	638	655	687	741	797	834	836	805	818	840	855	864	832	823	813	777	727
Z10ALL	SYS3	2097	717	1329	6	1	2	4	5	6	6	6	6	6	6	7	7	7	7	6	6	6	6	6	6	6	6	6
Z10ALL	SYS4	2097	717	1329	45	51	45	42	30	21	20	23	27	31	35	38	37	38	39	43	65	54	34	34	32	32	31	29
NOLIMIT	SYS5	2097	717	1329	8	3	4	6	7	8	8	8	8	8	8	9	9	9	9	23	35	47	8	8	8	8	8	8
NOLIMIT	SYS6	2097	717	1329	8	3	4	6	7	8	8	8	8	8	8	9	9	9	9	22	37	48	8	8	8	8	8	8
Z10ALL	TOTAL				844	796	774	772	743	720	727	731	763	809	869	927	959	956	926	967	1039	1070	978	946	926	908	873	833

Drilling down further still, and the mystery was solved...

Figure 5

For some reason, the SYS5 and SYS6 LPARs were not included in the Z10ALL group and were therefore not controlled by the group capacity limit. So, in the peak hour, they used about 95 MSUs, which, on top of the 975 used by the Z10ALL group, led to a total of 1070 MSUs being used.

4 Elementary my dear Watson!

The explanation was, as often happens, very simple.

By looking at the EPV Exceptions, they found an alert pointing to a wrong Group Capacity definition.

	GROUP CAPACITY CONFIGURATION - THU, 26 JAN 2012														
CEC	GROUP	SYSTEM	LPAR- NAME	CEC MSU	GROUP MSU	WEIGHT	DEF MSU	MIN ENT	MAX ENT	CAP	OLD Z/ OS	DED	WC=Y		
SER1	Z10ALL	SYS1	LPR1	1329	1010	136	0	137.4	1010	Ν	Ν	Ν	Ν		
SER1	Z10ALL	SYS2	LPR2	1329	1010	717	0	724.2	1010	Ν	Ν	Ν	Ν		
SER1	Z10ALL	SYS3	LPR3	1329	1010	5	9	5.1	9	Ν	Ν	Ν	Ν		
SER1	Z10ALL	SYS4	LPR4	1329	1010	70	126	70.7	126	Ν	Ν	Ν	Ν		
SER1	Z10ALL	SYS5	LPR5	1329	1010	36	0	36.4	1010	Y	Ν	Ν	Ν		
SER1	Z10ALL	SYS6	LPR6	1329	1010	36	0	36.4	1010	Y	N	N	Ν		

Figure 6

On January 26th, it was decided to hard cap SYS5 and SYS6 before running a new application performance test. Unfortunately, as explained in the WLM manual, when the limitations described in Section 2 above (Group Capacity Overview), are not fulfilled:

"All partitions which do not conform to these rules are not considered part of the group. WLM will dynamically remove such partitions from the group and manage the remaining partitions towards the group limit."

In all fairness to the customer, we have to say that these hardware capping limitations were not documented in either the z/OS 1.10 WLM manual, the above sentence, nor the z/OS WLM manual, prior to 1.10.

The description of the limitations was incomplete and is outlined below:

"WLM will only manage partitions with shared CPs and running on z/OS V1R8. All partitions which do not conform to this rule will not be considered as part of the group."

5 Conclusions

Group Capacity limit is a very powerful tool which is able to protect z/OS customers from unexpected and undesired software cost increases.

However, it is important to be aware that LPAR definitions have to comply to the group capacity rules and limitations.

In this paper we described a real-life situation where the lack of knowledge unwittingly caused an increase in the monthly peak of the 4-hour rolling average of about 60 MSUs.

This "oversight" led to extra costs - in this case - of around \$78,000.

SEGUS Inc is the North American distributor for EPV products

For more information regarding EPV for z/OS, please visit www.segus.com or call (800) 327-9650