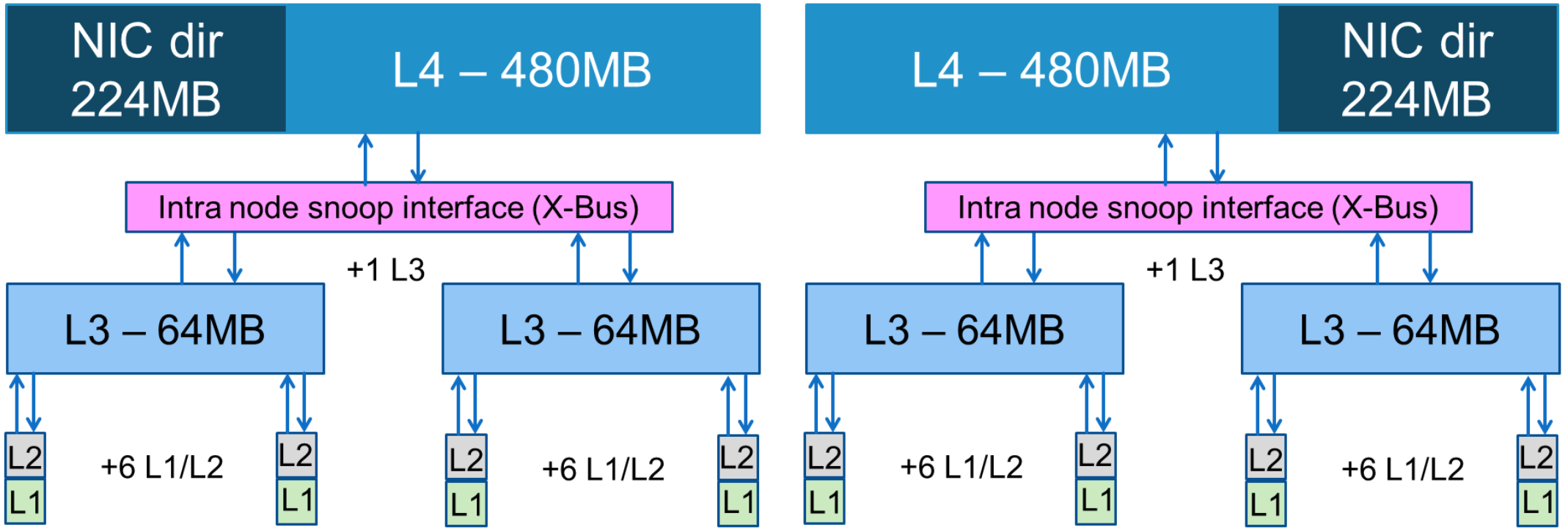
z14 Capacity Planning: theory and practice

Danilo Gipponi
EPV Technologies

Agenda

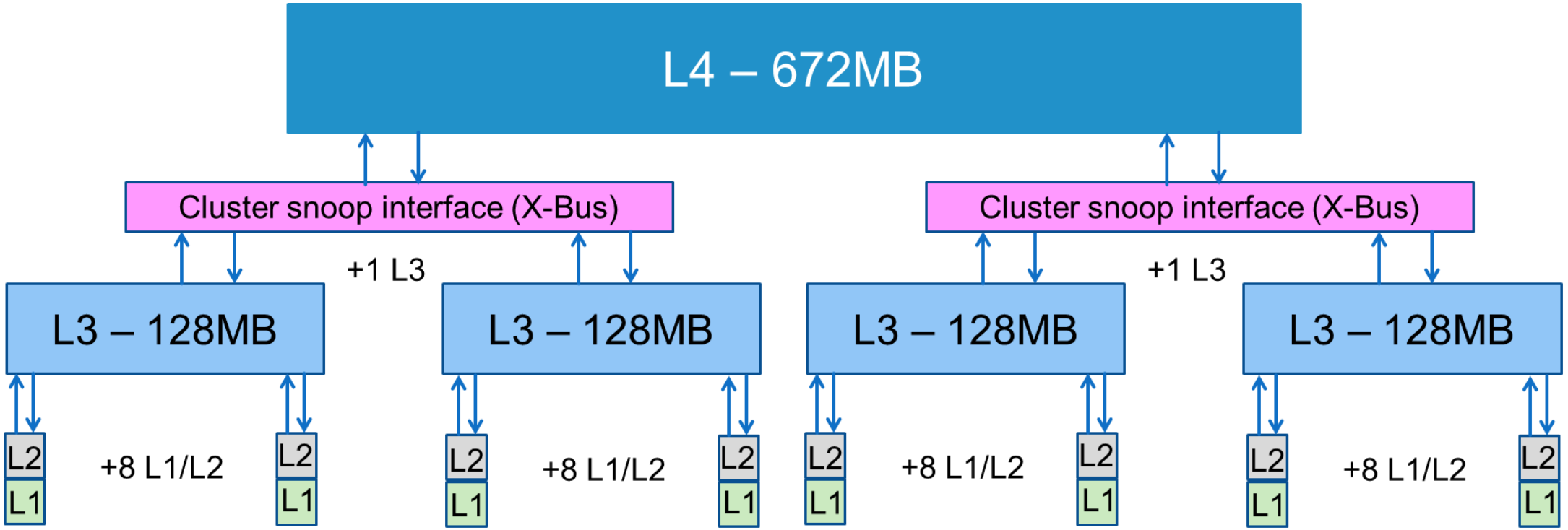
- z14 processor cache architecture
- New SMF 113 counters for z14
- Choosing the right benchmark
- Capacity planning and CPI
- Other MF Indexes
- User experiences
- Summary

z13 processor cache architecture



L2	Instructions	2MB
	Data	2MB
L1	Instructions	96K
	Data	128K

z14 processor cache architecture

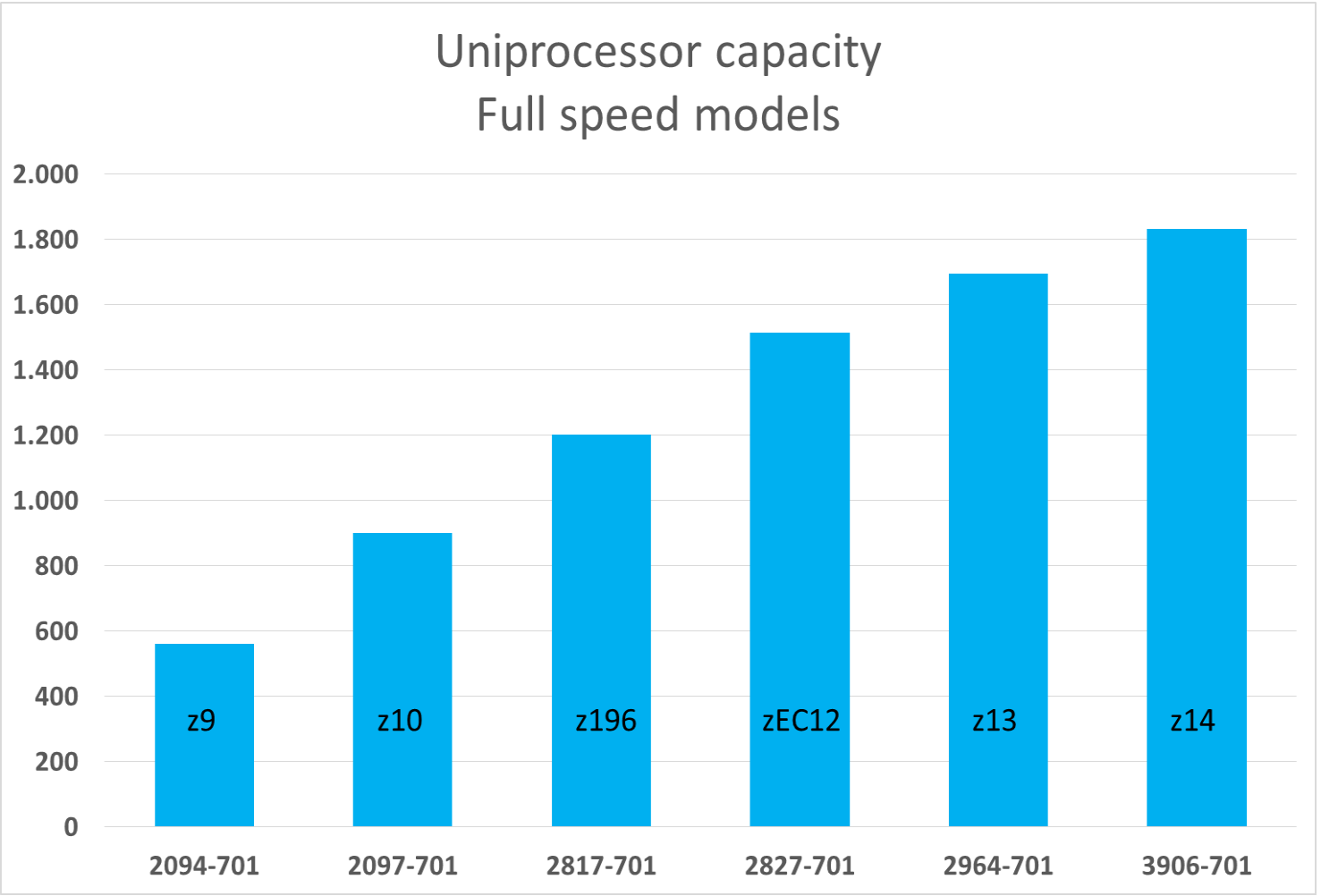


L2	Instructions	2MB
	Data	4MB
L1	Instructions	128K
	Data	128K

z14 processor cache architecture

- Advantages:
  - ✓Bigger L1 cache for instructions; from 96K to 128K
  - ✓Bigger L2 cache for data; from 2MB to 4MB
  - ✓Bigger L3 cache; from 64MB to 128MB
  - ✓Full inclusive L4 cache; no NIC anymore
  - ✓L4 cache common to clusters inside a drawer
  - ✓Point to point communication among drawers
- Disadvantages:
  - ✓Smaller L4 cache; from 960MB (480MB x 2) to 672MB

z14 processor cache architecture



Single CP capacity improvement:

- 61% from z9 to z10
- 33% from z10 to z196
- 26% from z196 to zEC12
- 12% from zEC12 to z13
- 8% from z13 to z14

EPV TECHNOLOGIES

New SMF 113 counters for z14

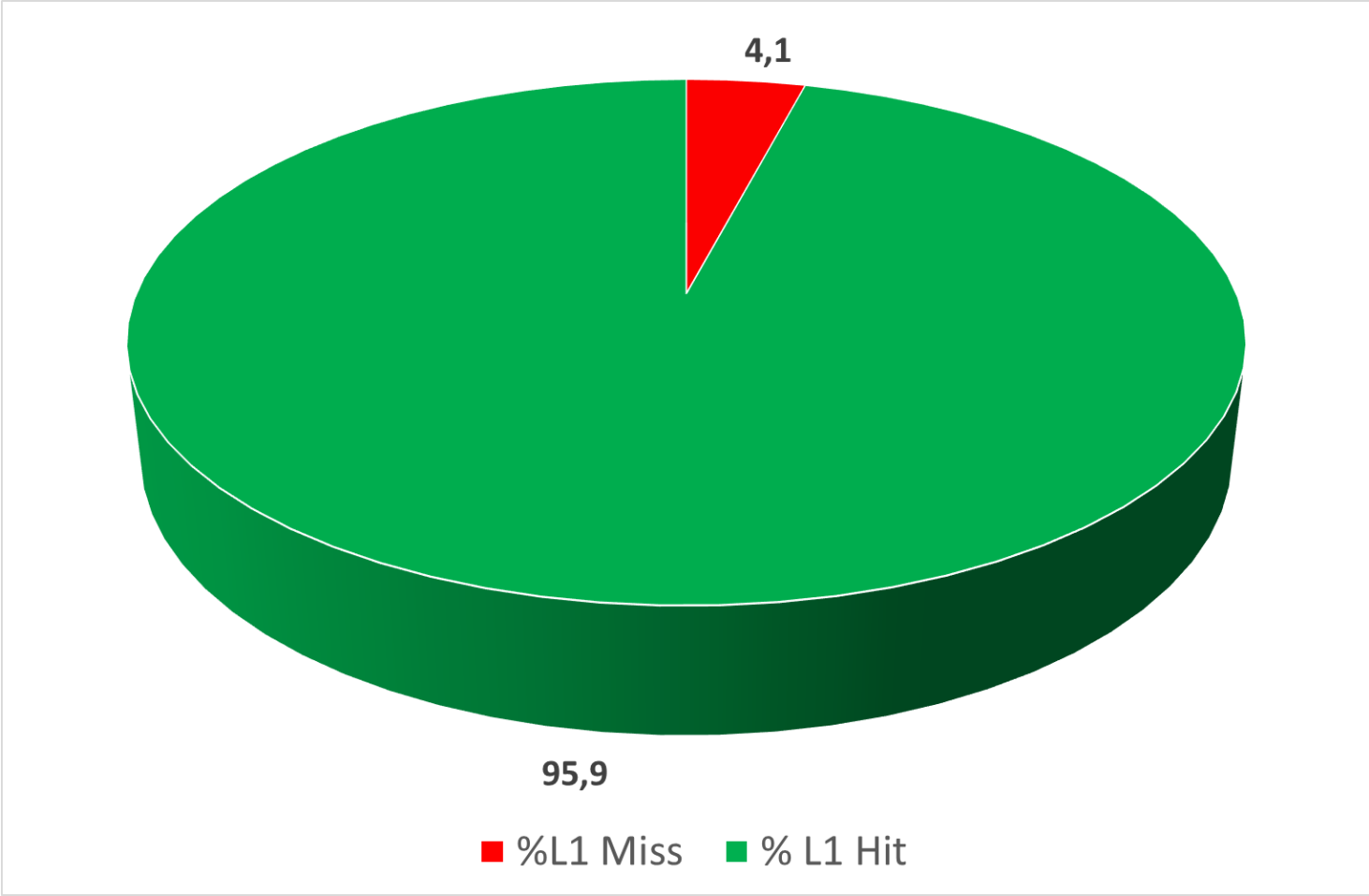
- Use SMF 113 subtype 1; values are de-accumulated; subtype 2 is frozen
- z14 can be identified in SMF 113 by using the SMF113_2_CTRVN2 variable (5=z14)
- Basic and Problem state counters have the same meaning as in z13
- As usual, extended counters are specific for z14; be careful
- TLB counters used in EPV are the same, even if the TLB architecture have also been changed; TLB1 has been integrated in the L1 caches

New SMF 113 counters for z14

- More details about the most useful counters in EPV white papers: “z14 Capacity Planning” Part 1, Part 2 and TLB addendum
- Complete extended counters information in “The CPU Measurement Facility Extended Counters Definition for z10, z196/z114, zEC12/zBC12, z13/z13s and z14” – SA23-2261-04

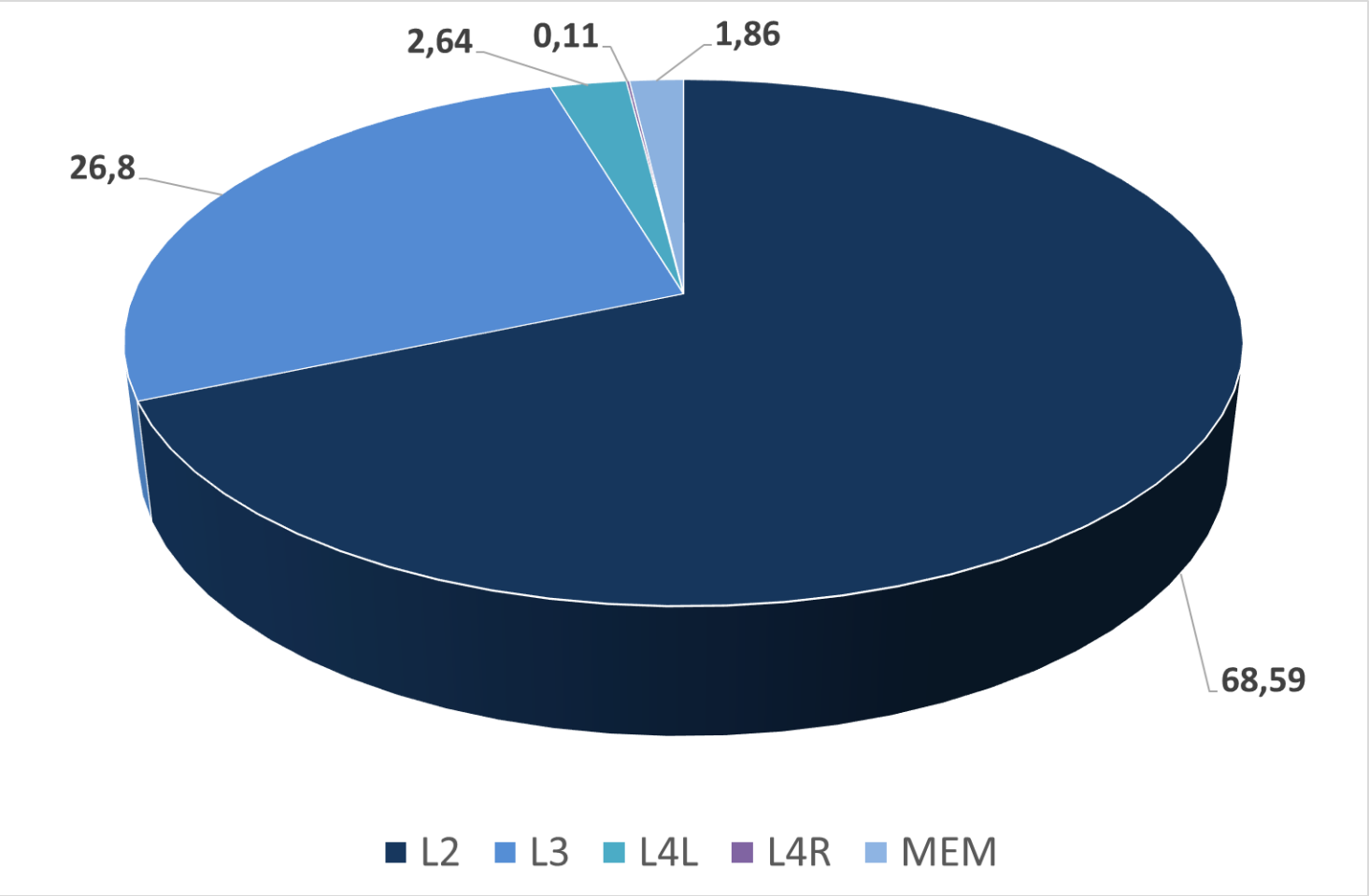
New SMF 113 counters for z14

- Real life example
- % L1 Miss in peak hours
- Based on basic counters



New SMF 113 counters for z14

- Real life example
- %Insourcing from other cache levels when a L1 Miss occurs
- Based on extended counters
- Used for RNI

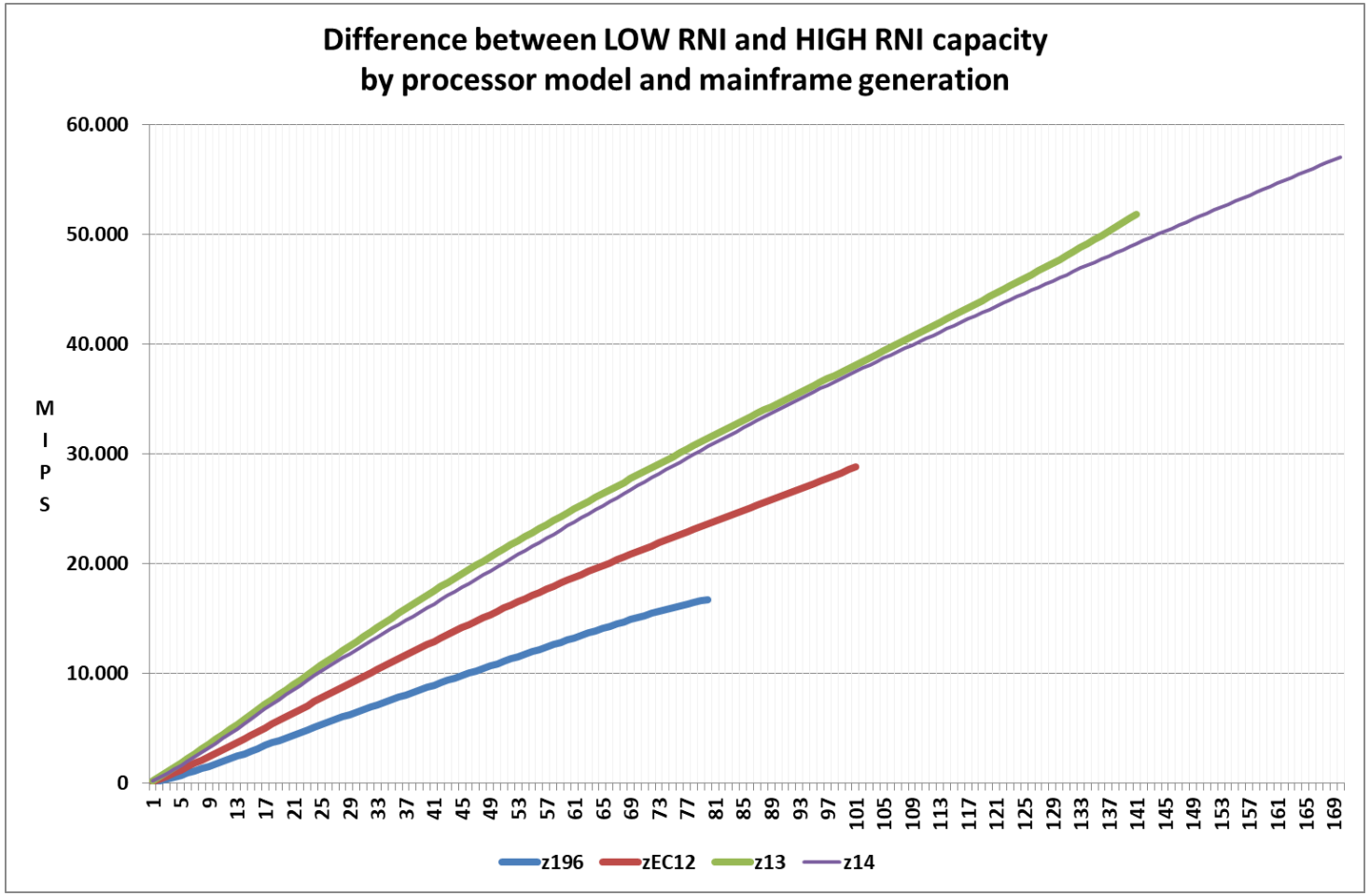


Choosing the right benchmark

Processor	#CP	PCI	MSU	Low	Average	High	Low MIPS	Avg MIPS	High MIPS
3906-701	1	1832	227	3,37	3,27	3,04	1.886,5	1.830,5	1.701,8
3906-702	2	3464	427	6,55	6,19	5,61	3.666,6	3.465,1	3.140,4
3906-703	3	5050	620	9,66	9,02	8,08	5.407,6	5.049,3	4.523,1
3906-704	4	6590	808	12,68	11,77	10,45	7.098,2	6.588,8	5.849,8
3906-705	5	8082	990	15,64	14,44	12,74	8.755,1	8.083,4	7.131,8
3906-706	6	9528	1162	18,55	17,02	14,96	10.384,1	9.527,7	8.374,5

EPV TECHNOLOGIES

Choosing the right benchmark



Choosing the right benchmark

%L1 Miss	RNI	Benchmark
< 3%	>= 0,75	AVG RNI
< 3%	< 0,75	LOW RNI
3% to 6%	> 1,00	HIGH RNI
3% to 6%	0,60 to 1,00	AVG RNI
3% to 6%	< 0,60	LOW RNI
> 6%	>= 0,75	HIGH RNI
> 6%	< 0,75	AVG RNI

Choosing the right benchmark

- %L1M formula as usual:

$$((B2 + B4) / B1) * 100$$

- New RNI formula for z14:

$$\mathbf{2.4} \times (\mathbf{0.4} \times \%L3 + \mathbf{1.5} \times \%L4L + \mathbf{3.2} \times \%L4R + \mathbf{7.0} \times \%MEM) / 100$$

where %L* is the percentage of L1 miss insourced from each cache level

EPV TECHNOLOGIES

Capacity planning and CPI

- CPI can be easily calculated as the ratio between B0 (cycles used) and B1 (instructions performed)
- CPI can also be split in two components:
  - ✓Finite CPI (cycles needed because of the L1 cache size limitation)
  - ✓Infinite CPI (cycles due to instructions complexity needed even with an inifinite L1 cache)
- These calculations are also described in EPV white papers

Capacity planning and CPI

- Special care is needed when CPI is used to compare two different machines
- Let's suppose that CPI in the peak hours is 3,0 on a zEC12
- After moving to z13 the CPI is now 2,7
- What is the performance improvement with the new machine?
- It seems to be 10%. Right?

Wrong: the answer is about 1%

- When doing such comparisons, CPI has to be weighted with processor Ghz that in zEC12 is 5,5 and in z13 "only" 5,0

Capacity planning and CPI

- Let's calculate million (M) inx/sec

$$zEC12\ M\text{-}inx/sec = \frac{5.500.000.000\ cycle/sec}{3{,}0\ cycle/inx} = 1.833\ M\text{-}inx/sec$$

$$z13\ M\text{-}inx/sec = \frac{5.000.000.000\ cycle/sec}{2{,}7\ cycle/inx} = 1.851\ M\text{-}inx/sec$$

- Difference is about 18 M-inx/sec that is about 1% improvement

Capacity planning and CPI

- You can get the same result by normalizing the zEC12 CPI using the z13 Ghz

*zEC12 norm to z13 = 3,0 cycl/inx / 5,5 Ghx * 5,0 Ghz = 2,73 cycle/inx*

- So the z13 benefit (CPI real reduction in our example) is

  (2,73 cycle/inx – 2,70 cycle/inx) / 2,73 cycle/inx = 0,03 / 2,73 = 1%

- z14 runs at 5,2 Ghz

EPV TECHNOLOGIES

Other MF Indexes (z13 example)

- Average number of cycles to serve a L1M (SCXL1M)

CEC1 CPU SCXL1M - 8-12,15-17											
		SYSTEMS									
DATE	DAY	SYS1	SYS2	SYS3	SYS4	SYS5	SYS6	SYS7	SYS8	SYS9	SYS10
06/10/2017	Fri	65,6	41,6	65,9	65,9	134,1	157,9	84,4	88,6	41,6	74,1
05/10/2017	Thu	70,2	44,9	65,9	69,2	131,1	155,6	81,8	87,7	40,5	75,4
04/10/2017	Wed	72,8	54,3	67,7	69,1	140,0	162,0	86,4	91,1	41,4	75,9
03/10/2017	Tue	69,0	46,1	61,8	68,3	130,7	157,3	82,8	87,0	37,3	73,5
02/10/2017	Mon	65,8	51,4	56,0	63,2	134,4	154,7	85,2	83,4	34,0	71,5

Other MF Indexes (z13 example)

- Average number of CPU cycles used for a TLB miss (CPU TLB1M)

CEC1 CPU TLB1M - 8-12,15-17											
		SYSTEMS									
DATE	DAY	SYS1	SYS2	SYS3	SYS4	SYS5	SYS6	SYS7	SYS8	SYS9	SYS10
06/10/2017	Fri	69,3	55,9	76,2	83,3	99,0	102,2	83,5	84,6	63,3	66,4
05/10/2017	Thu	70,9	52,5	76,1	82,3	97,1	101,6	81,7	84,4	63,7	66,0
04/10/2017	Wed	71,0	56,5	77,5	81,1	101,5	104,6	84,3	86,7	63,3	67,2
03/10/2017	Tue	69,8	55,2	73,6	80,1	97,2	101,8	81,5	83,1	65,2	64,6
02/10/2017	Mon	67,6	53,7	75,5	82,1	98,0	100,1	79,1	81,3	62,9	66,6

Other MF Indexes (z13 example)

- Percentage of CPU cycles used for TLB misses (%CPU TLB1M)

CEC1 %CPU TLB1M - 8-12,15-17											
		SYSTEMS									
DATE	DAY	SYS1	SYS2	SYS3	SYS4	SYS5	SYS6	SYS7	SYS8	SYS9	SYS10
06/10/2017	Fri	4,1	4,2	5,7	6,4	5,7	5,7	5,9	5,9	5,1	6,5
05/10/2017	Thu	4,0	4,1	5,6	6,3	5,5	5,6	5,8	5,9	5,1	6,5
04/10/2017	Wed	4,1	4,1	5,7	6,0	5,7	5,7	5,6	5,9	5,0	6,6
03/10/2017	Tue	3,8	3,9	5,5	5,7	5,5	5,7	5,3	5,5	4,6	6,5
02/10/2017	Mon	3,7	3,9	4,6	5,2	5,5	5,6	5,6	5,6	4,3	6,3

Other MF Indexes (z13 example)

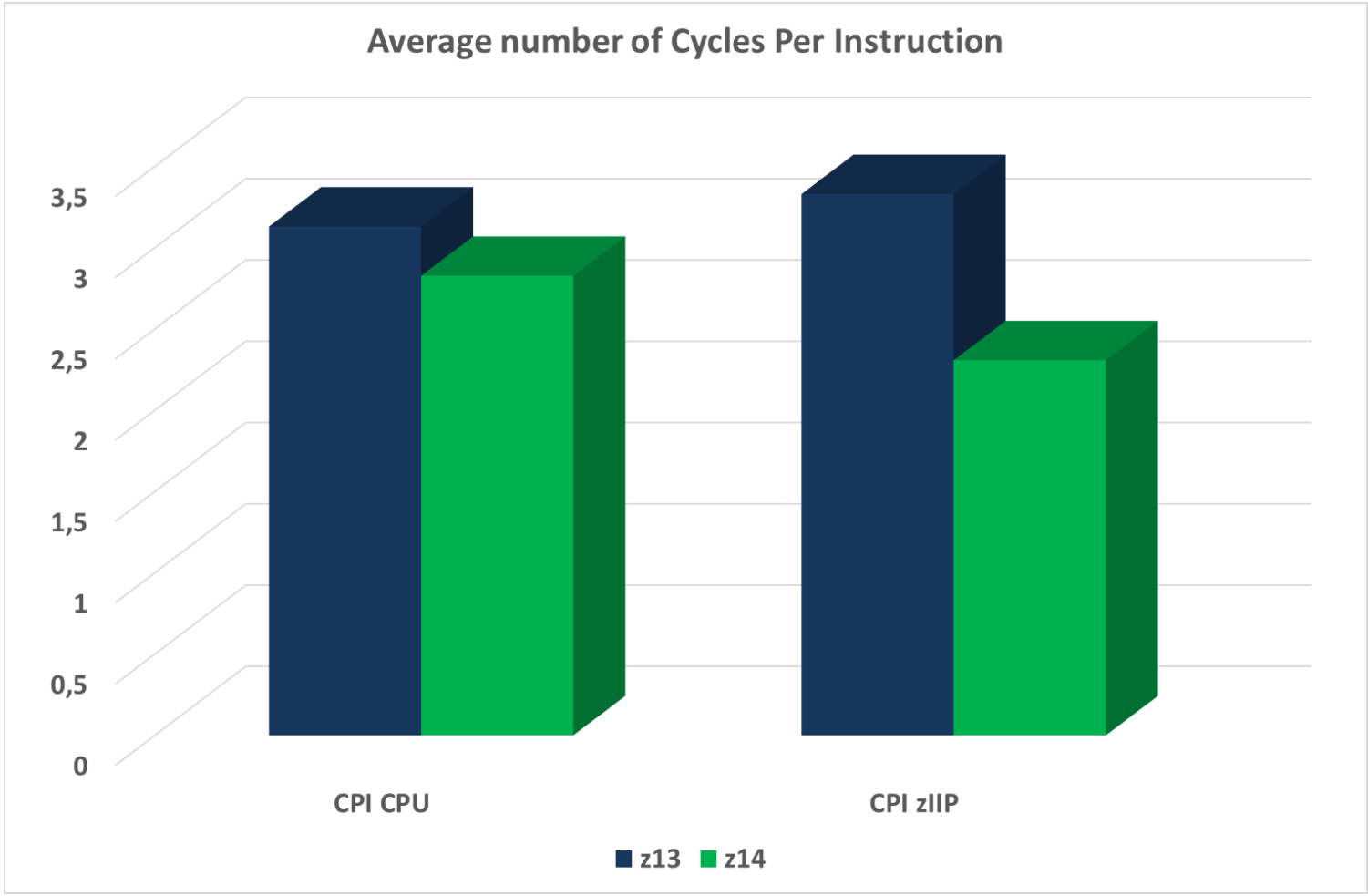
- Cache Residency time

CEC1 CACHE RESIDENCY TIME - 8-12,15-17 - MSEC					
DATE	DAY	L1	L2	L3	L4
06/10/2017	Fri	0,05	3,08	53,82	490,00
05/10/2017	Thu	0,05	3,03	52,97	476,15
04/10/2017	Wed	0,05	2,81	48,57	454,83
03/10/2017	Tue	0,05	3,09	54,34	482,88
02/10/2017	Mon	0,05	3,07	53,60	454,40

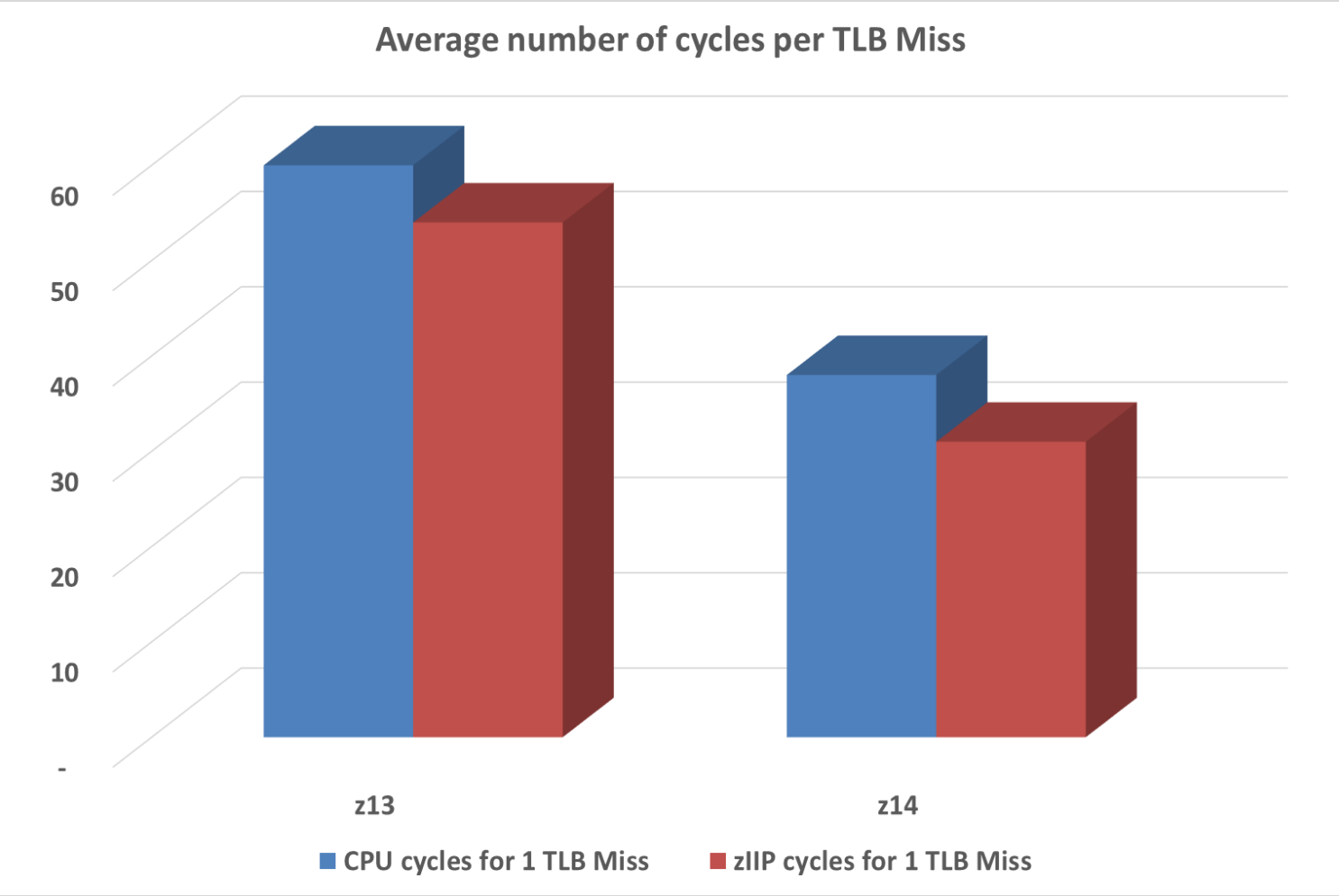
User experiences: case 1

- Customer migration from z13 to z14
- Only some days before and after the hardware upgrade have been collected so these results have to be considered just as an indication to be confirmed with more data
- Overall impression is that z14 performs significantly better than z13

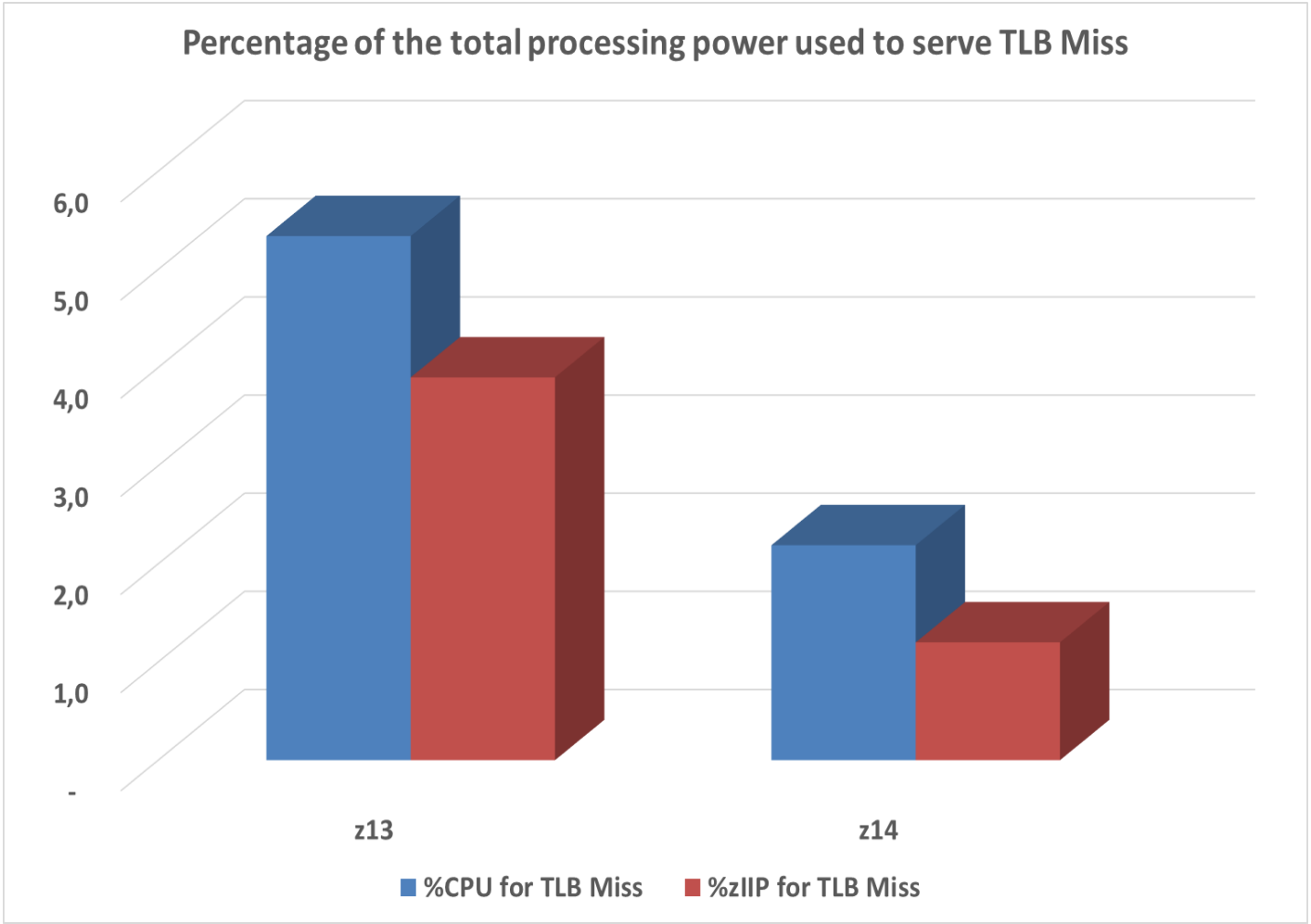
User experiences: case 1



User experiences: case 1



User experiences: case 1



User experiences: case 1

- Estimated residency time by cache level in milliseconds

	L1	L2	L3	L4
z13	0,020	1,740	84,850	921,550
z14	0,040	3,885	266,885	628,520

User experiences: case 2

- Customer migration from zEC12 to z14
- Small customer
- Italian branch outsourced to the mother company
- Overall impression is that z14 performs better than zEC12

User experiences: case 2

- CEC CPUs reduced from 6 to 5
- Same number of zIIP, ICF and IFL
- Small increase of MSU and MIPS (AVG RNI benchmark used)

CECS - Fri, 20 Apr 2018										
CEC	TYPE	HW	MDL	MSU	CPU MIPS	CPUS	ICFS	AAPS	IIPS	IFLS
	2827	H43	706	957	7.809,00	6	2	0	4	1
CECS - Mon, 23 Apr 2018										
CEC	TYPE	HW	MDL	MSU	CPU MIPS	CPUS	ICFS	AAPS	IIPS	IFLS
	3906	M02	705	990	8.083,40	5	2	0	4	1

User experiences: case 2

- Same number of logical processors on each LPAR
- Memory slightly increased

SYSTEMS - Fri, 20 Apr 2018															
CEC	SYSTEM	LPARNAME	SYSNAME	SYSPLEX	OS LEVEL	GMTOFF	HDISP	CPUS	AAPS	IIPS	CORE	MT CPU	MT IIP	MEMORY	LFAREA
	SYSG	P22S	SYSG	GANPLX00	ZV020200	2	Y	1	0	1	N	1	1	2.048	
	SYSP	P21S	SYSP	GANPLX00	ZV020200	2	Y	2	0	3	N	1	1	65.536	16.384
	SYST	P23S	SYST	GANPLX00	ZV020200	2	Y	1	0	1	N	1	1	16.384	1.024

SYSTEMS - Mon, 23 Apr 2018															
CEC	SYSTEM	LPARNAME	SYSNAME	SYSPLEX	OS LEVEL	GMTOFF	HDISP	CPUS	AAPS	IIPS	CORE	MT CPU	MT IIP	MEMORY	LFAREA
	SYSG	P22S	SYSG	GANPLX00	ZV020200	2	Y	1	0	1	N	1	1	8.192	
	SYSP	P21S	SYSP	GANPLX00	ZV020200	2	Y	2	0	3	N	1	1	73.728	16.384
	SYST	P23S	SYST	GANPLX00	ZV020200	2	Y	1	0	1	N	1	1	24.576	1.024

User experiences: case 2

zEC12		L1M	RNI
20/04/2018	Fri	3,56	0,72
19/04/2018	Thu	3,37	0,72
18/04/2018	Wed	3,45	0,74
17/04/2018	Tue	3,74	0,72
16/04/2018	Mon	3,64	0,68
13/04/2018	Fri	3,51	0,71
12/04/2018	Thu	3,51	0,71
11/04/2018	Wed	3,51	0,71
10/04/2018	Tue	3,50	0,71
09/04/2018	Mon	3,50	0,71

%L1 Miss	RNI	Benchmark
< 3%	>= 0,75	AVG RNI
< 3%	< 0,75	LOW RNI
3% to 6%	> 1,00	HIGH RNI
3% to 6%	0,60 to 1,00	AVG RNI
3% to 6%	< 0,60	LOW RNI
> 6%	>= 0,75	HIGH RNI
> 6%	< 0,75	AVG RNI

User experiences: case 2

z14		L1M	RNI
04/05/2018	Fri	2,98	0,72
03/05/2018	Thu	2,40	0,94
02/05/2018	Wed	2,55	0,80
30/04/2018	Mon	2,50	0,75
27/04/2018	Fri	2,66	0,76
26/04/2018	Thu	2,82	0,67
24/04/2018	Tue	2,67	0,78
23/04/2018	Mon	2,45	0,81

%L1 Miss	RNI	Benchmark
< 3%	>= 0,75	AVG RNI
< 3%	< 0,75	LOW RNI
3% to 6%	> 1,00	HIGH RNI
3% to 6%	0,60 to 1,00	AVG RNI
3% to 6%	< 0,60	LOW RNI
> 6%	>= 0,75	HIGH RNI
> 6%	< 0,75	AVG RNI

User experiences: case 2

SYSTEM	DATE	DAY	%L2	%L3	%L4L	%L4R	%MEML	%MEMR
SYSP	04/05/2018	Fri	68,02	25,39	4,79	0,03	1,77	-
SYSP	03/05/2018	Thu	66,23	25,55	5,19	0,03	3,01	-
SYSP	02/05/2018	Wed	67,62	24,95	5,17	0,03	2,24	-
SYSP	30/04/2018	Mon	69,55	24,25	3,95	0,03	2,23	-
SYSP	27/04/2018	Fri	68,84	24,54	4,46	0,03	2,13	-
SYSP	26/04/2018	Thu	72,63	21,35	4,09	0,02	1,91	-
SYSP	24/04/2018	Tue	67,72	25,11	5,00	0,03	2,14	-
SYSP	23/04/2018	Mon	68,03	24,45	5,20	0,03	2,30	-
SYSP	20/04/2018	Fri	71,34	20,51	6,26	0,02	0,98	0,89
SYSP	19/04/2018	Thu	70,58	21,00	6,55	0,02	0,94	0,90
SYSP	18/04/2018	Wed	70,45	21,05	6,57	0,02	0,98	0,93
SYSP	17/04/2018	Tue	69,23	22,65	6,32	0,02	0,93	0,85
SYSP	16/04/2018	Mon	72,57	19,78	5,87	0,02	0,90	0,86
SYSP	13/04/2018	Fri	71,53	20,52	6,05	0,02	0,95	0,92
SYSP	12/04/2018	Thu	71,54	20,52	6,05	0,02	0,95	0,92
SYSP	11/04/2018	Wed	71,55	20,51	6,05	0,02	0,95	0,92
SYSP	10/04/2018	Tue	71,58	20,48	6,04	0,02	0,96	0,92
SYSP	09/04/2018	Mon	71,61	20,46	6,03	0,02	0,96	0,93

- L1 Miss insourcing
- Relative percentages

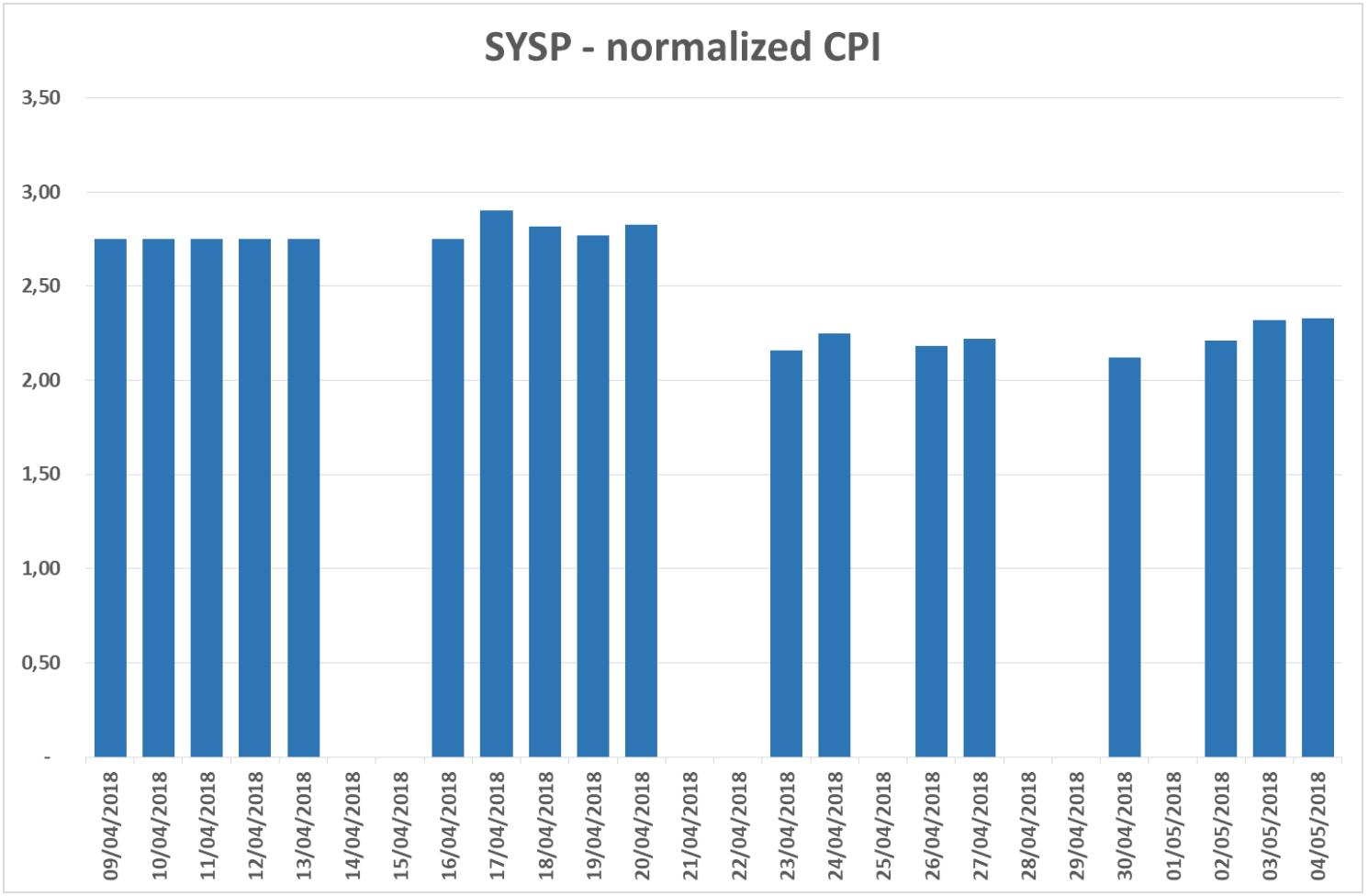
EPV TECHNOLOGIES

User experiences: case 2

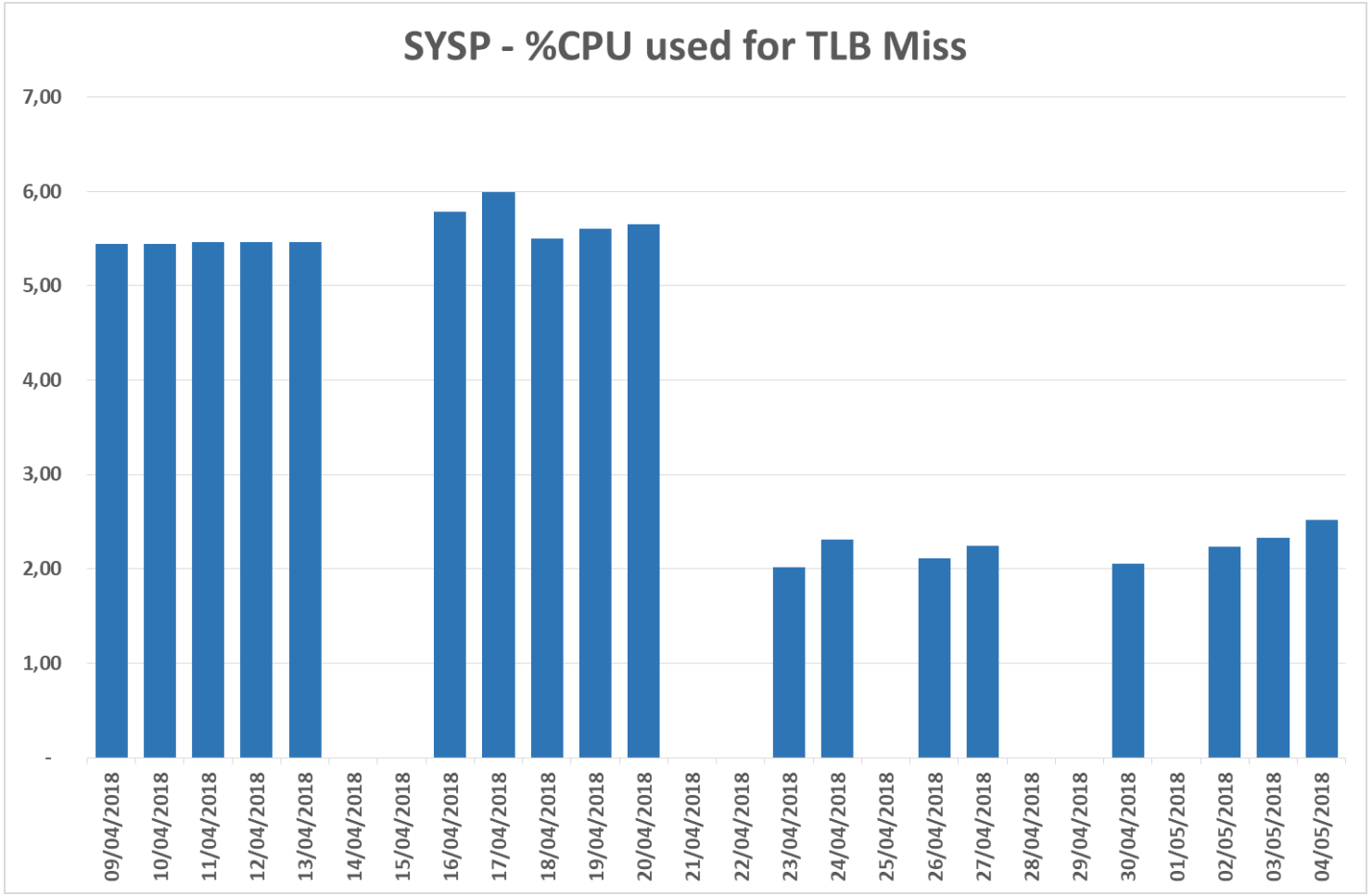
- Relative weight is about the same
- But HiperDispatch polarization changed
- 2 medium instead of 1 high and 1 medium

zEC12												
SYSTEM	LPARNAME	CEC MSU	GROUP MSU	WEIGHT	%WEIGHT	Target CPUs	VH	VM	VL	DEF MSU	MIN ENT	MAX ENT
SYSP	P21S	957	293	260	27,2%	1,63	1,00	1,00		0	260,9	293
z14												
SYSTEM	LPARNAME	CEC MSU	GROUP MSU	WEIGHT	%WEIGHT	Target CPUs	VH	VM	VL	DEF MSU	MIN ENT	MAX ENT
SYSP	P21S	990	293	270	27,3%	1,36		2,00		0	261,1	293

User experiences: case 2



User experiences: case 2



Summary

- Very small increase of z14 uni-processor speed
- Size of cache levels increased (not for the L4)
- New interesting optimizations of cache processor architecture
- SMF 113 data collection and reporting absolutely needed to measure system performance
- General impression about z14 is positive

Questions?

danilo.gipponi@epvtech.com

epv.support@epvtech.com

www.epvtech.com

EPV TECHNOLOGIES