



Advanced Technical Skills (ATS) North America

CPU MF – 2012 Update and WSC Experiences

SHARE - Session 11600

August 8, 2012

John Burg

jpburg@us.ibm.com

IBM
Washington Systems Center



Agenda

- **New Dawn in System z Capacity Planning**
 - What and Why
- **Validation and Enablement**
- **CPU MF Metrics**
 - Basic
 - RNI
 - z196 / z114 RNI Update July 2012
- **Data Profiles**
- **HiperDispatch Considerations**
- **Summary**

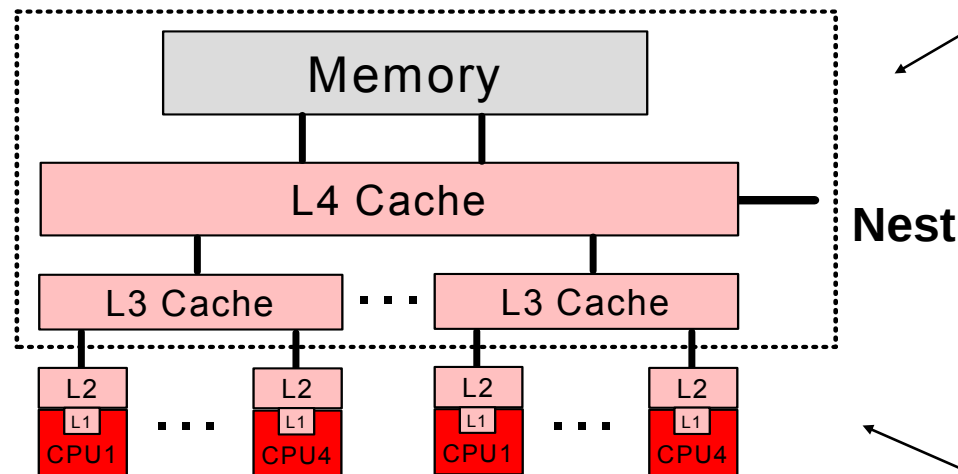
New Day Dawning in System z Capacity Planning

Processor Design

- CPU
- Memory Hierarchy (Nest)

Hypervisor (PR/SM)

- Amount of virtualization



Operating System

- Virtualization at address space level

Workload Characteristics

- Instructions
- Dispatch Profile
- I/O Rate

Introduction to LSPR

- **A set of representative SCP/workload environments**
 - SCPs: z/OS, z/VM, and Linux on System z
 - Workload categories: Low ←Relative Nest Intensity→ High
 - Current LSPR workload categories: Low, Average, High
 - zPCR extends published categories
 - Low-Avg
 - Avg-High
 - A methodology focused on processor capacity
 - No significant external constraints
 - Equivalent (reasonably high, e.g. $\geq 90\%$) processor utilization
- **A metric to communicate the results**
 - ITR: Internal Throughput Rate
 - Transactions or Jobs per processor busy second
- **Information stored on the web**
 - <https://www.ibm.com/servers/resourcelink/lib03060.nsf/pages/lspindex?OpenDocument>

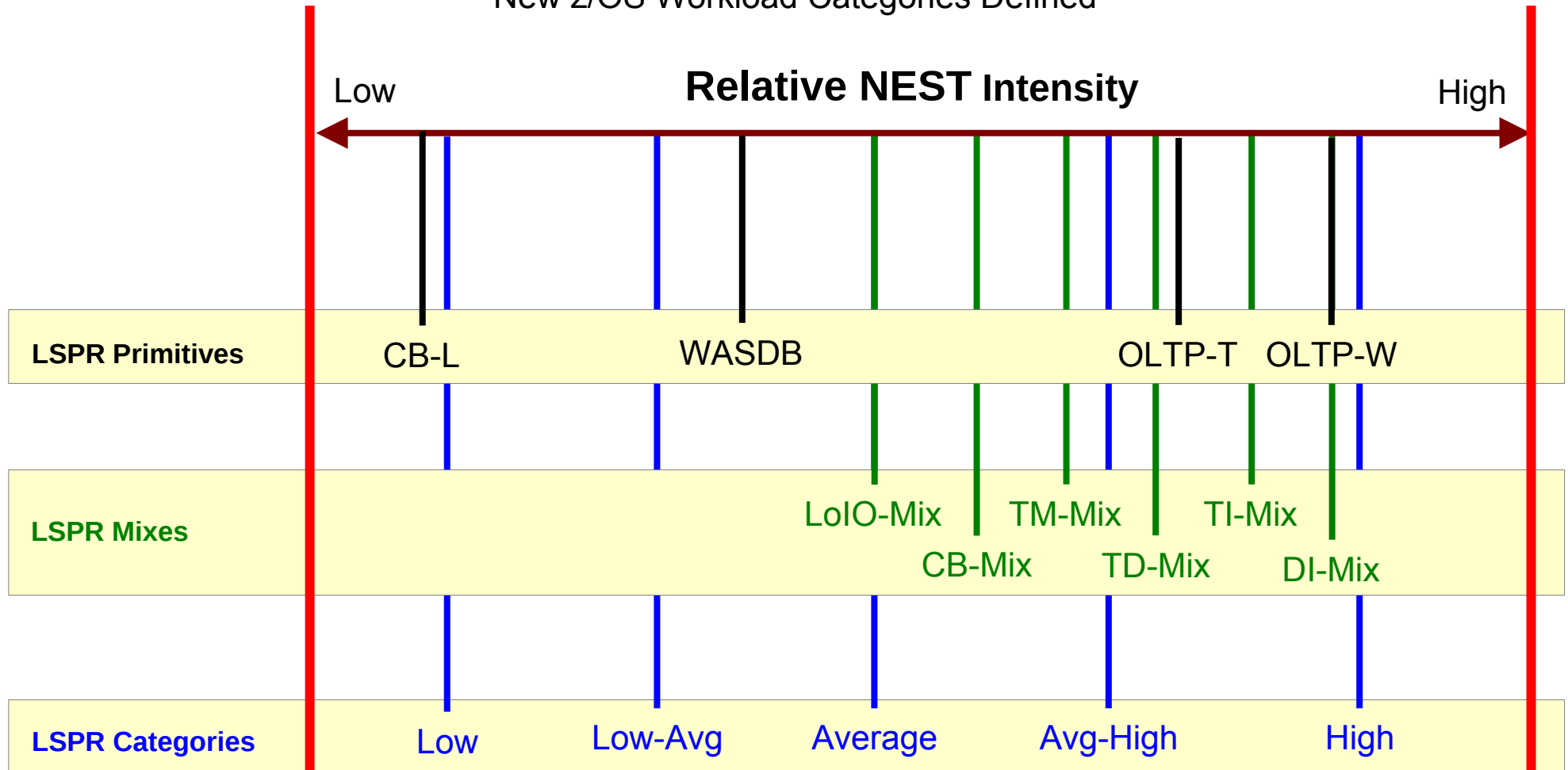
LSPR Workload Categories

- Various combinations of workload primitives are measured on which the new workload categories are based
 - Applications include CICS, DB2, IMS, OSAM, VSAM, WebSphere, COBOL, utilities
- **Low** (relative nest intensity)
 - Workload curve representing light use of the memory hierarchy
 - Similar to past high scaling workload primitives
- **Average** (relative nest intensity)
 - Workload curve expected to represent the majority of customer workloads
 - Similar to the past LoIO-mix curve
- **High** (relative nest intensity)
 - Workload curve representing heavy use of the memory hierarchy
 - Similar to the past DI-mix curve
- zPCR extends published categories
 - **Low-Avg**
 - 50% Low and 50% Average
 - **Avg-High**
 - 50% Average and 50% High

zPCR Workload Characterization for z/OS

“Scope of Work” Definition Change

New z/OS Workload Categories Defined



Use zPCR's Workload Selection Assistant to choose appropriate workload category

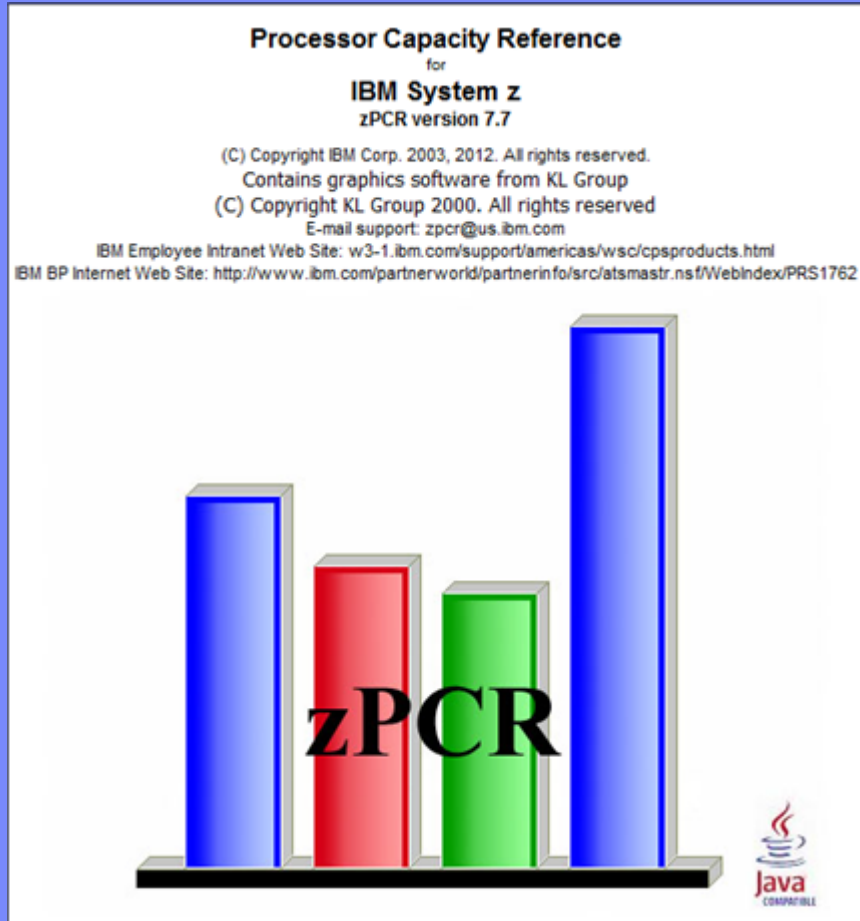
Automated with EDF input into zPCR

Note: Workload selection is automated in zCP3000

CPU Measurement Facility

- **Introduced in z10 and later processors**
- **Facility that provides hardware instrumentation data for production systems**
- **Two Major components**
 - Counters
 - Cache and memory hierarchy information
 - SCPs supported include z/OS and z/VM
 - Sampling
 - Instruction time-in-CSECT
- **New z/OS HIS started task**
 - Gathered on an LPAR basis
 - Writes SMF 113 records
- **New z/VM Monitor Records**
 - Gathered on an LPAR basis – all guests are aggregated
 - Writes new Domain 5 (Processor) Record 13 (CPU MF Counters) records
- **Minimal overhead**

New Hardware Capabilities to Size z/OS Workloads



LSPR BENCHMARKS

Manual Input

RMF Report

CP3KEXTR

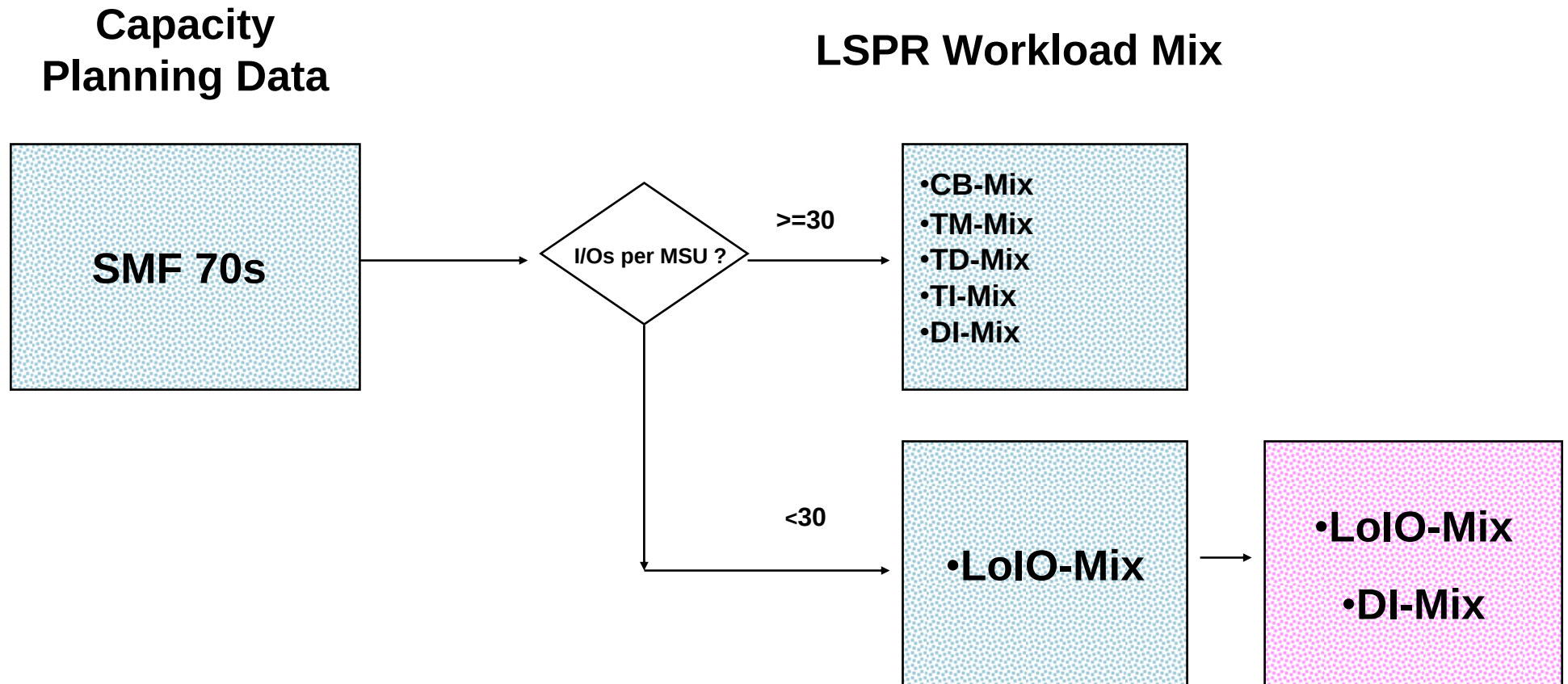
SMF 70
•LPAR Data

SMF 113
•Counters cache /
memory hierarchy

Importance of using CPU MF Counters

- **New CPU MF Counters provide better information to do more successful capacity planning**
- **Same data used to validate the LSPR workloads can now be obtained from production systems**
 - Matches your production workload to the LSPR workloads
 - zPCR automatically processes CPU MF data to provide a match
 - Based on Relative Nest Intensity (RNI)
- **CPU MF Counters also useful for performance analysis**

Challenge to Use SMF to Select a LSPR Workload Mix



SMF 113s Provide Better LSPR Workload Selection

Capacity Planning Data

SMF 70s and 113s



**Relative Nest Intensity
(RNI) &
L1 Miss % (L1MP)**



LSPR Workload Category

- Low
- Low-Average
- Average
- Average-High
- High

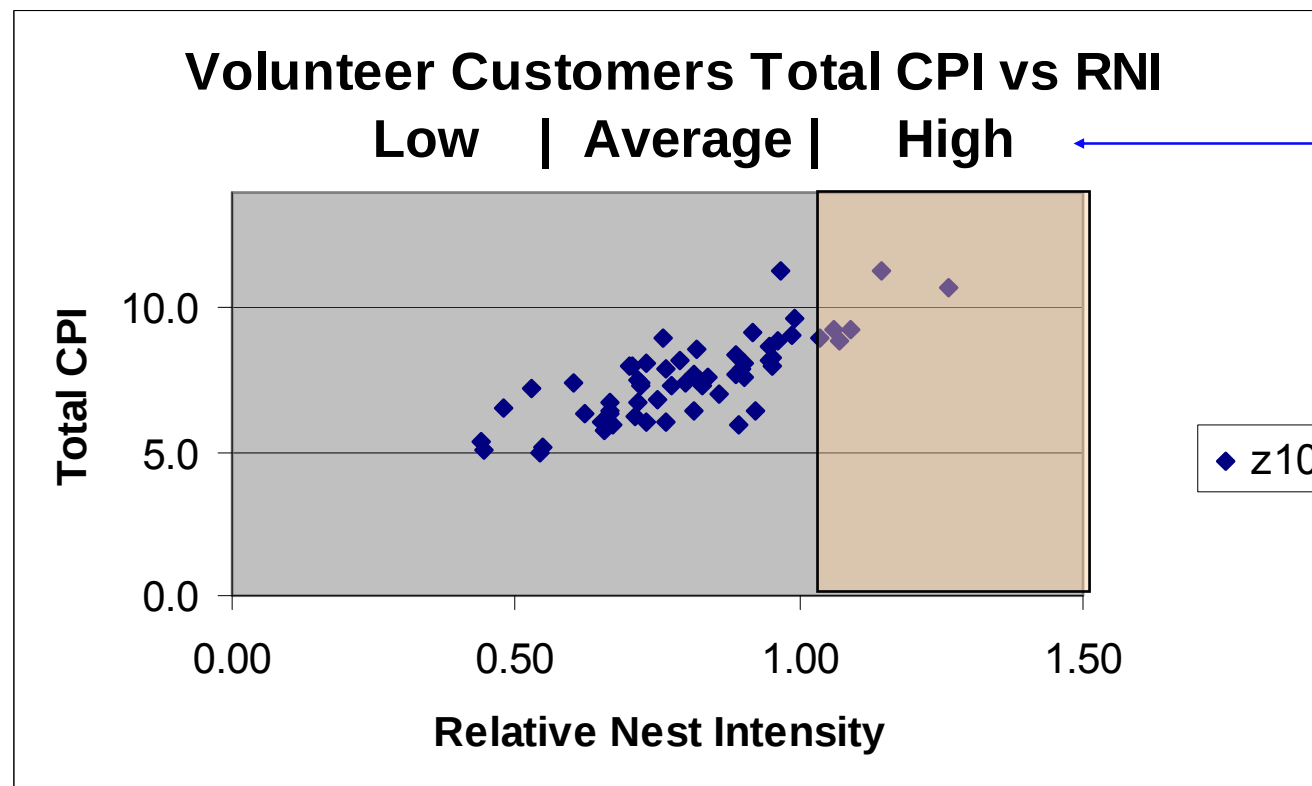
Validation and Enablement Details

Workload Category Validation Process

- **Analyzed customer data**
- **Initial z10 Study**
 - 100 z10 customer LPARs
 - SMF 70s and SMF 113s
 - Built relationship between performance and Nest
 - Built new metric to more precisely match workloads
 - Relative Nest Intensity (RNI)
 - LSPR workload categories defined as Low, Average, and High RNI
- **Validated RNI metrics with review of z10 to z196 Migrations**
 - 75 z10 to z196 customer LPAR migrations
 - Validated RNI based workload match

CPU MF

z10 Customer Workload Characterization Summary



3) Created new
[LSPR Workload
Categories](#)

1) Customer [CPI](#)
measurements

2) Created new [RNI](#) metric

Requirements to Enable CPU MF Counters

- **Processor requirements**

- z10, z196, or z114

- **z/OS requirements**

- z/OS 1.10 or higher
 - With APAR OA30486

- **z/VM requirements**

- z/VM at 5.4 or higher
 - With APAR VM64961

z/OS Steps to Enable CPU MF Counters

- **1 - Configure the processor to collect CPU MF**

- ___ Update the LPAR Security Tabs, can be done dynamically

- **2 - Set up HIS and z/OS to collect CPU MF**

- ___ Set up HIS Proc

- ___ Set up OMVS Directory - required

- ___ Collect SMF 113s via SMFPRMxx

- **3 - Collect CPU MF COUNTERs**

- ___ Start HIS

- ___ Modify HIS: "F HIS,B,TT='Text',PATH='/his/',CTRONLY,CTR=(B,E),SI=SYNC"

- Recommend to start HIS, Modify for Counters, and continuously run

SMF 113s Space Requirements

- The SMF 113 record puts minimal pressure on SMF
 - 452 bytes for each logical processor per interval
- Example below is from 3 z196s processors
 - 713, 716 and 718
 - 10 Systems
 - 5 Days, 24 hours
- SMF 113s were 1.2% of the space compared to SMF 70s & 72s

RECORD TYPE	RECORDS READ	PERCENT OF TOTAL	AVG. RECORD LENGTH	MIN. RECORD LENGTH	MAX. RECORD LENGTH	RECORDS WRITTEN	Total Size (with AVG. Record Size)	% Total Size (with AVG. Record Size)
70	14,250	1.8%	14,236	640	32,736	14,250	202,865,850	15.1%
72	744,014	93.5%	1,516	1,104	20,316	744,014	1,128,252,590	83.7%
113	37,098	4.7%	452	452	452	37,098	16,768,296	1.2%
TOTAL	795,362	100.0%	1,695	18	32,736	795,362	1,347,886,736	100.0%

Operations – Display Command

F HIS,B,TT='BE Counters',PATH='/his/',CTRONLY,CTR=(B,E),SI=SYNC

D HIS

RESPONSE=SYSD

HIS015I 10.15.54 DISPLAY HIS 286

HIS 0025 ACTIVE

COMMAND: MODIFY HIS,B,TT='BE Counters',PATH='/his/',CTRONLY,CTR=(B,E),
SI=SYNC

START TIME: 2012/04/12 10:15:45

END TIME: ----/--/-- --:--:--

COMPLETION STATUS: -----

FILE PREFIX: SYSHIS20120412.101545.

COUNTER VERSION NUMBER 1: 1 COUNTER VERSION NUMBER 2: 2

COMMAND PARAMETER VALUES USED:

TITLE= BE Counters

PATH= /his/

COUNTER SET= BASIC, EXTENDED

DURATION= NOLIMIT

CTRONLY

DATALOSS= IGNORE

STATECHANGE= SAVE

SMFINTVAL= SYNC

Use CPU MF Counters for Performance Analysis

- **Counters can be used as a secondary source to:**
 - Supplement current performance data from SMF, RMF, DB2, CICS, etc.
 - Help understand why performance may have changed

- **Some examples of usage include:**
 - HiperDispatch Impact
 - Configuration changes (Additional LPARs)
 - 1 MB Page implementation
 - Application Changes (e.g. CICS Threadsafe Vs QR)
 - Estimating Utilization Effect for capacity planning
 - z196 GHz change in Power Saving Mode
 - Crypto CPACF usage

Metrics

z196 versus z10 hardware comparison

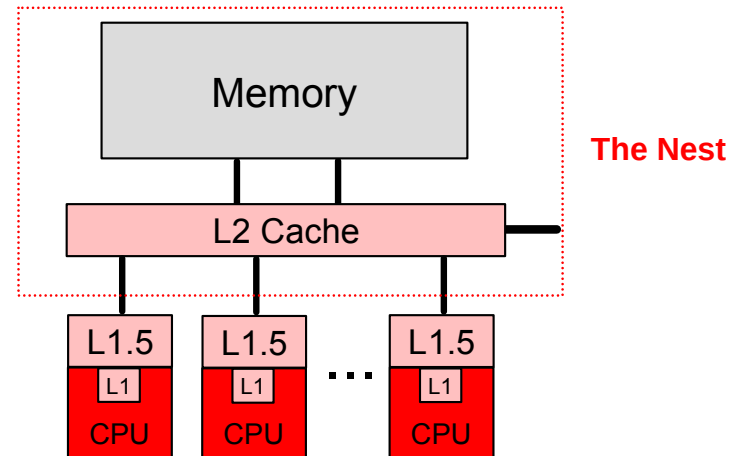
■ z10 EC

▶ CPU

- 4.4 GHz

▶ Caches

- L1 private 64k i, 128k d
- L1.5 private 3 MB
- L2 shared 48 MB / book
- book interconnect: star



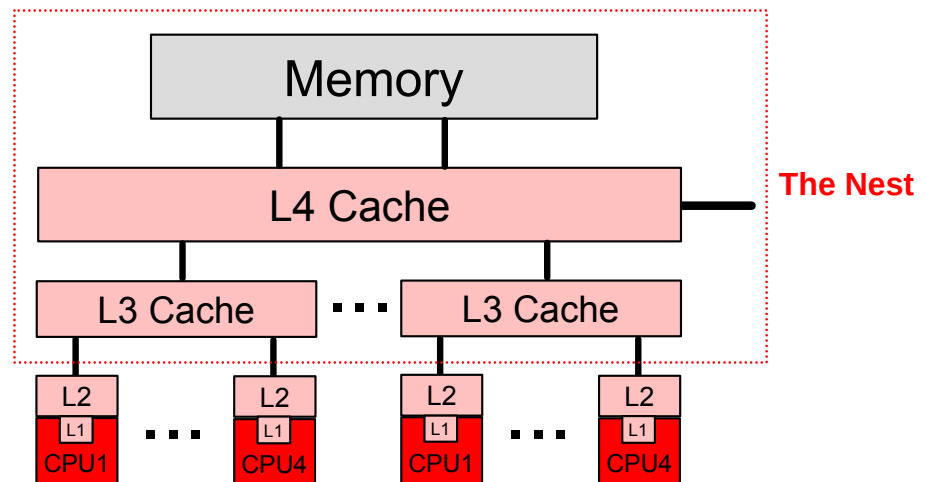
■ z196

▶ CPU

- 5.2 GHz
- Out-Of-Order execution

▶ Caches

- L1 private 64k i, 128k d
- L2 private 1.5 MB
- L3 shared 24 MB / chip
- L4 shared 192 MB / book
- book interconnect: star



z114 hardware comparison

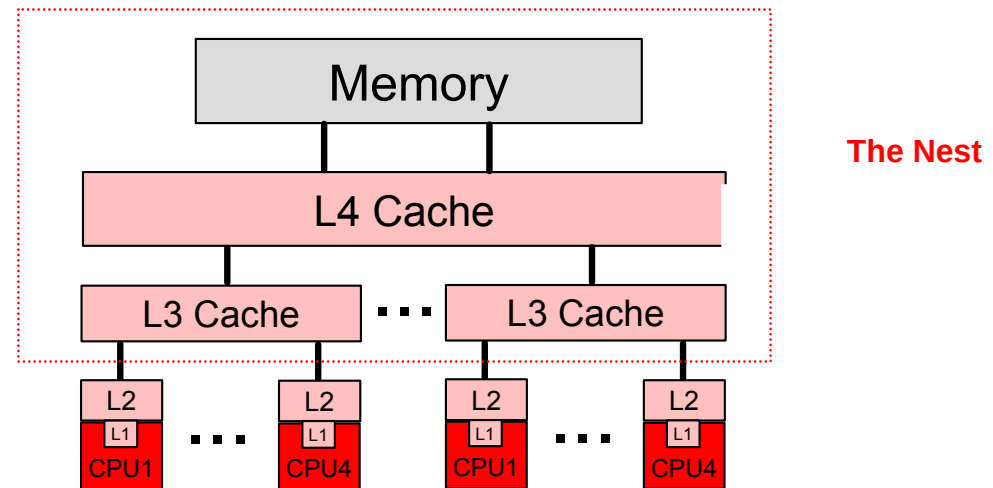
■ z114

▶ CPU

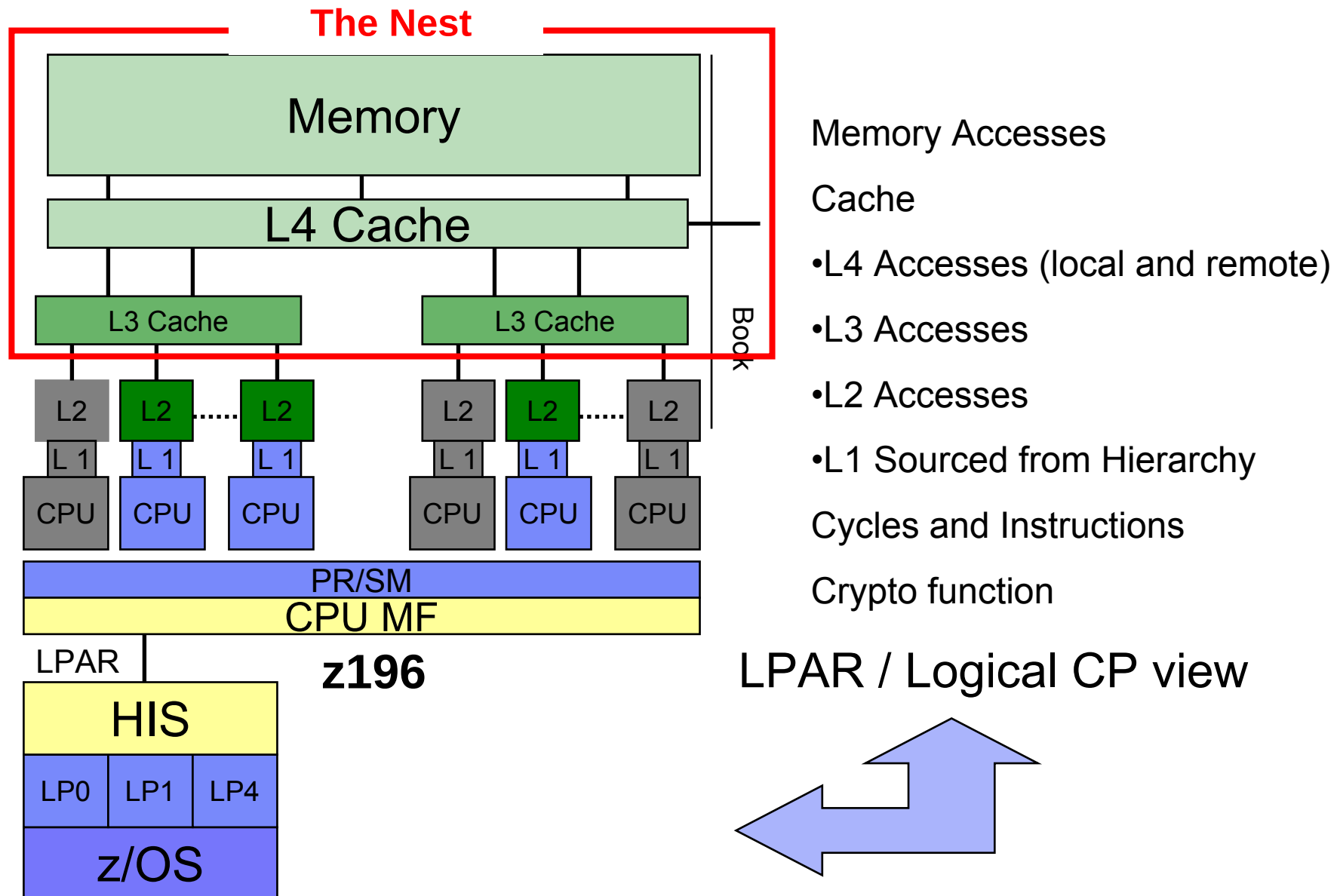
- 3.8 GHz
- Out-Of-Order execution

▶ Caches

- L1 private 64k i, 128k d
- L2 private 1.5 MB
- L3 shared 12 MB / chip
- L4 shared 96 MB / book
 - 24 MB to each core



z196 CPU MF Cache / Memory Hierarchy Sourcing



CPU MF Basic Performance Metrics:

CPI	Prb State	L1MP	L15P / L2P	L3P	L2LP / L4LP	L2RP / L4RP	MEMP	LPARCPU
-----	-----------	------	---------------	-----	----------------	----------------	------	---------

CPI – Cycles per Instruction

PRB STATE - % Problem State

L1MP – Level 1 Miss Per 100 instructions

L15P / L2P – % sourced from L1.5 or L2 cache

L3P – % sourced from L3 cache

L2LP / L4LP – % sourced from Level 2 (or L4) Local cache (on same book)

L2RP / L4RP – % sourced from Level 2 (or L4) Remote cache (on different book)

MEMP - % sourced from Memory

LPARCPU - APPL% (GCPs, zAAPs, zIIPs) captured and uncaptured


 Workload Characterization
 L1 Sourcing from cache/memory hierarchy

Workload Capacity Performance

- Instruction Complexity (Micro Processor Design)
 - Many design alternatives
 - Cycle time (GHz), instruction architecture, pipeline, superscalar, Out-Of-Order, branch prediction and more
 - Workload effect
 - May be different with each processor design
 - **Once established for a workload on a processor, doesn't change very much**

Workload Capacity Performance

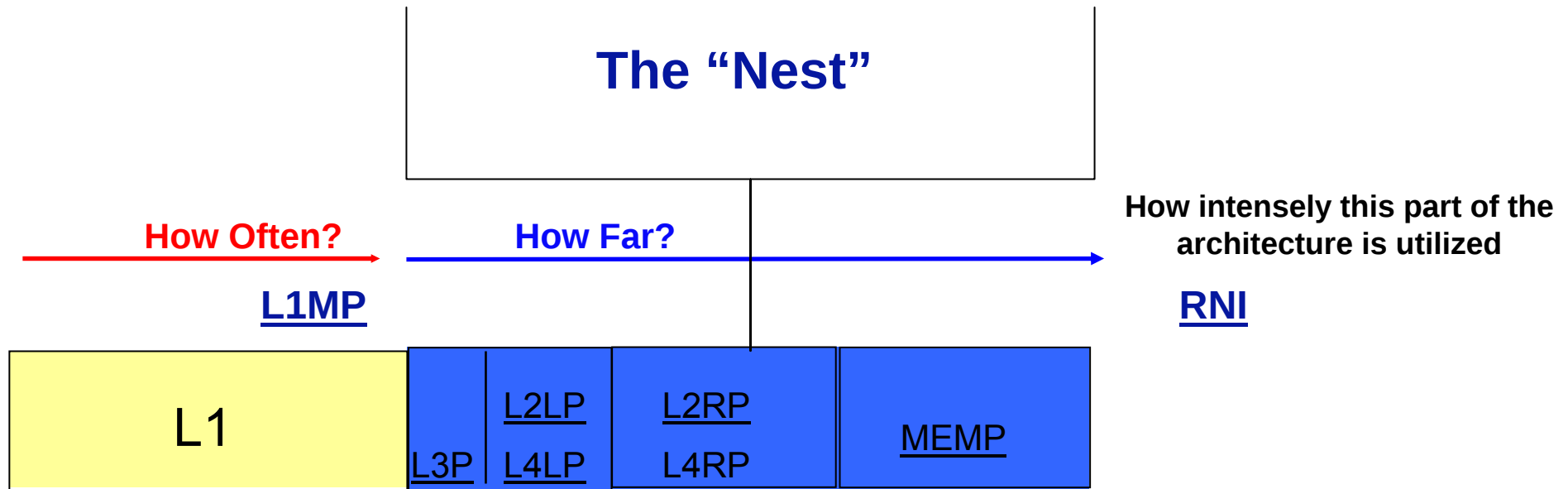
■ Memory Hierarchy or “Nest”

- Many design alternatives
 - Cache (levels, size, private, shared, latency, MESI protocol), controller, data buses
- Workload effect
 - Quite variable
 - **Sensitive to many factors: locality of reference, dispatch rate, IO rate, competition with other applications and/or LPARs, and more**
 - Net effect of these factors represented in “Relative Nest Intensity”
- Relative Nest Intensity (RNI)
 - **Activity beyond private-on-chip cache(s) is the most sensitive area**
 - **Reflects distribution and latency of sourcing from shared caches and memory**
 - Level 1 cache miss per 100 instructions (L1MP) also important
 - Data for calculation available from CPU MF (SMF 113) starting with z10

Relative Nest Intensity (RNI) Metric

- Reflects the distribution and latency of sourcing from shared caches and memory
 - For z10 EC and BC $RNI = (1.0 * L2LP + 2.4 * L2RP + 7.5 * MEMP) / 100$
 - For z196 / z114 $*RNI = 1.67 * (0.4 * L3P + 1.0 * L4LP + 2.4 * L4RP + 7.5 * MEMP) / 100$

Relative Nest Intensity



Microprocessor Design

Memory Hierarchy or Nest

*z196 / z114 RNI Changed July 2012

Note these Formulas may change in the future

z196 / z114 RNI Update – July 2012

- **z196 / z114 RNI updated**
 - Was $RNI = 1.60 * (0.4 * L3P + 1.0 * L4LP + 2.4 * L4RP + 7.5 * MEMP) / 100$
 - Now $RNI = \underline{1.67} * (0.4 * L3P + 1.0 * L4LP + 2.4 * L4RP + 7.5 * MEMP) / 100$
 - Refinement based on customer measurements and LSPR workloads
 - No change to z10 RNI calculation
- **zPCR 7.9b and CP3000 updated as of 7/31/12**
- See slides 51 to 55 for all metric definitions

z196 / z114 RNI Changed July 2012

Note these Formulas may change in the future

RNI-based LSPR Workload Decision Table

L1MP	RNI	LSPR Workload Match
<3%	≥ 0.75	AVERAGE
	< 0.75	LOW
3% to 6%	> 1.0	HIGH
	0.6 to 1.0	AVERAGE
	< 0.6	LOW
$> 6\%$	≥ 0.75	HIGH
	< 0.75	AVERAGE

Notes: applies to z10, z196 and z114 CPU MF data
table may change based on feedback

CPU MF Additional Performance Metrics:

CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L15P / L2P	L3P	L2LP / L4LP	L2RP / L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Machine Type	LSPR Wkld
-----	-----------	------------------------	-------------------	---------------	------	---------------	-----	----------------	----------------	------	-----------------------	---------	---------	-----------------	--------------

Est Instr Cmplx CPI – Estimated Instruction Complexity CPI

Est Finite CPI - Estimated Finite CPI

Est SCPL1M – Estimated Sourcing Cycles per L1 Miss Per 100 instructions

Eff GHz – Effective Gigahertz

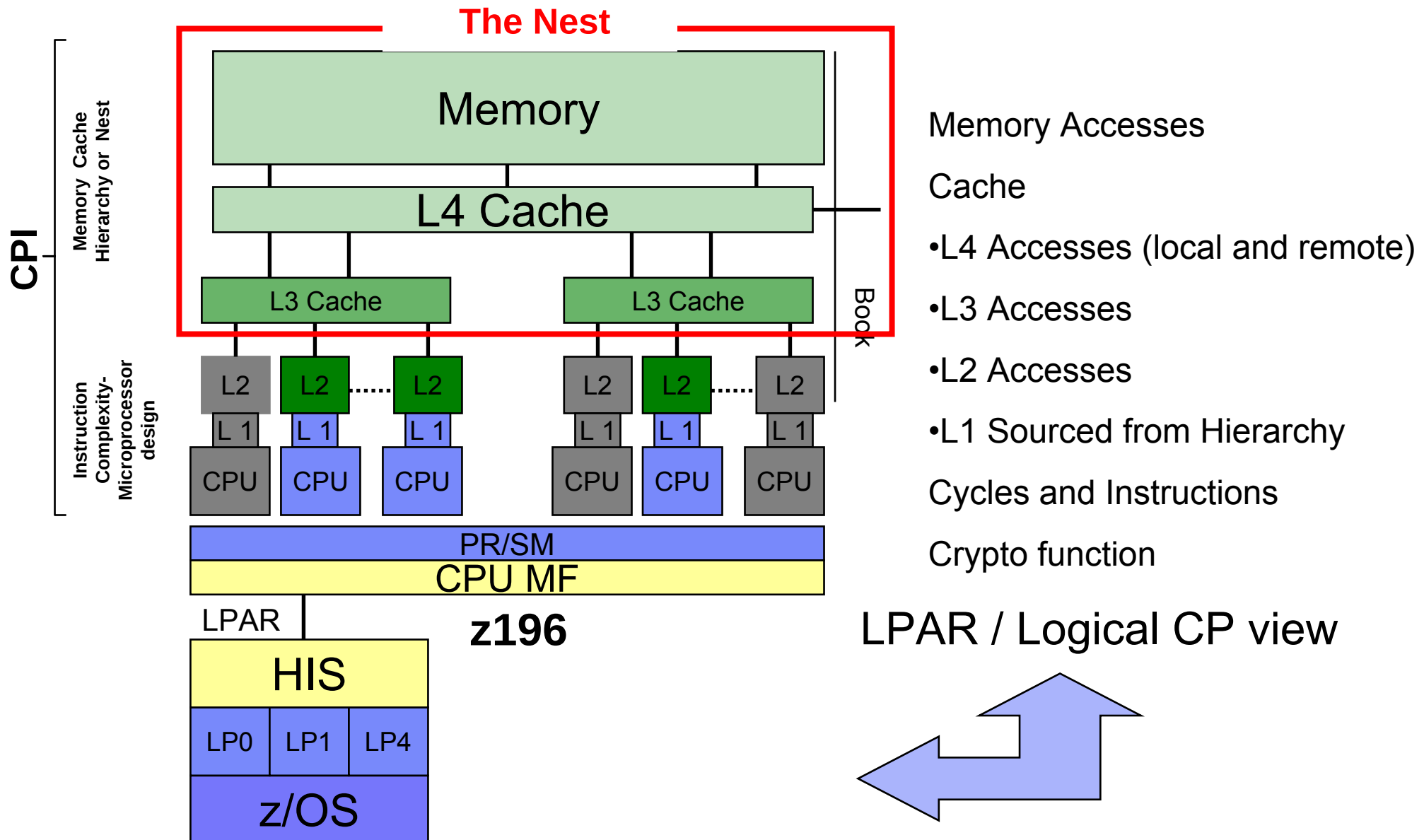
Machine Type – Machine Type (e.g. z10, z196, z114)

LSPR Wkld – LSPR Workload match based on L1MP and RNI

See slide 54 for z196 / z114 Est Finite CPI & Est SCPL1M updates – July 2012

Workload Characterization
L1 Sourcing from cache/memory hierarchy

z196 CPU MF Cache / Memory Hierarchy Sourcing



Data Profiles

Profiles

- **40 Total LPARs**
 - 14 z10 ECs / z10 BCs
 - 26 z196s / z114s

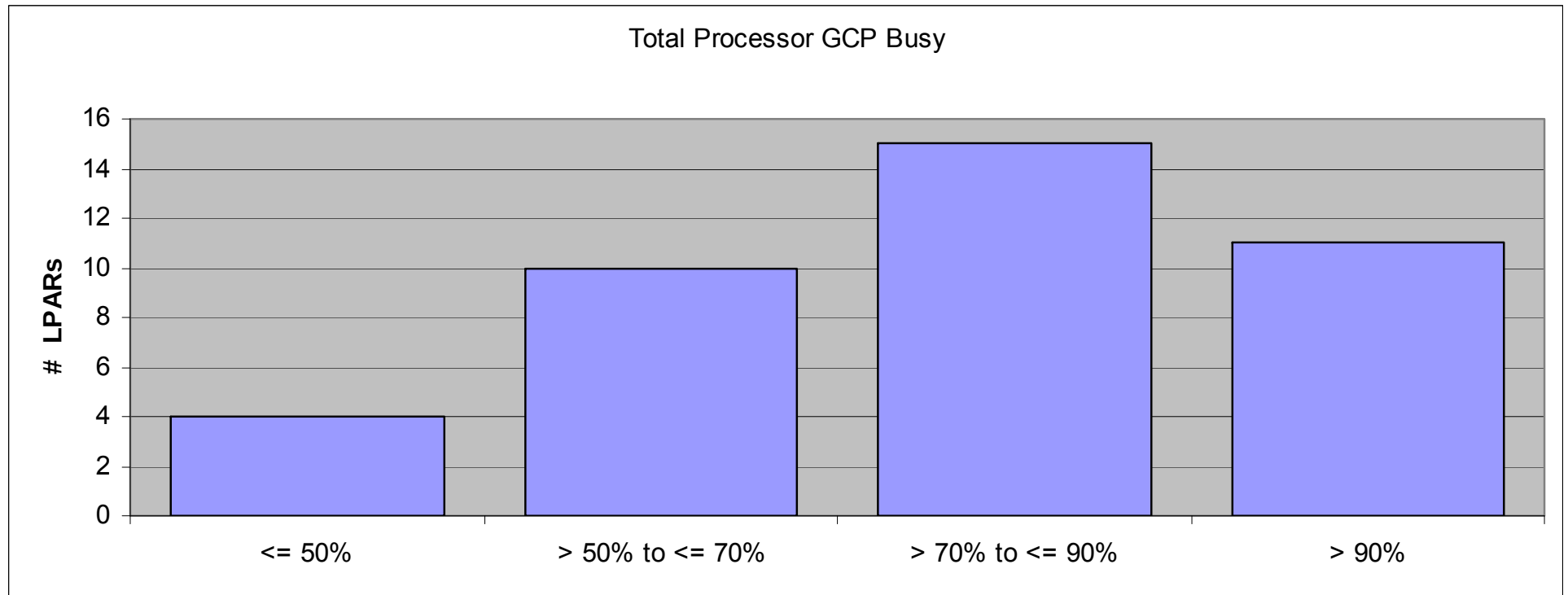
- **HiperDispatch**
 - 28 Yes
 - 12 No

- **33 LPARs utilized zIIPs and/or zAAPs**

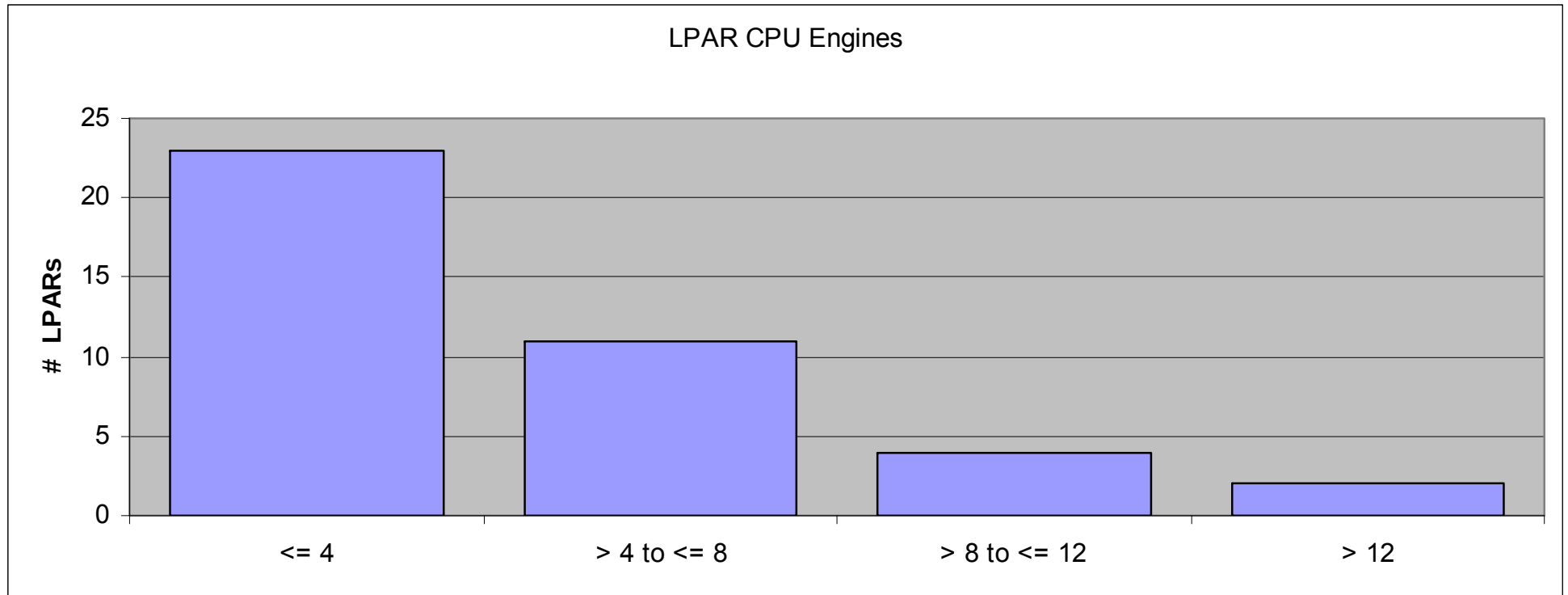
Customer Data from 2012

Completed June 2012

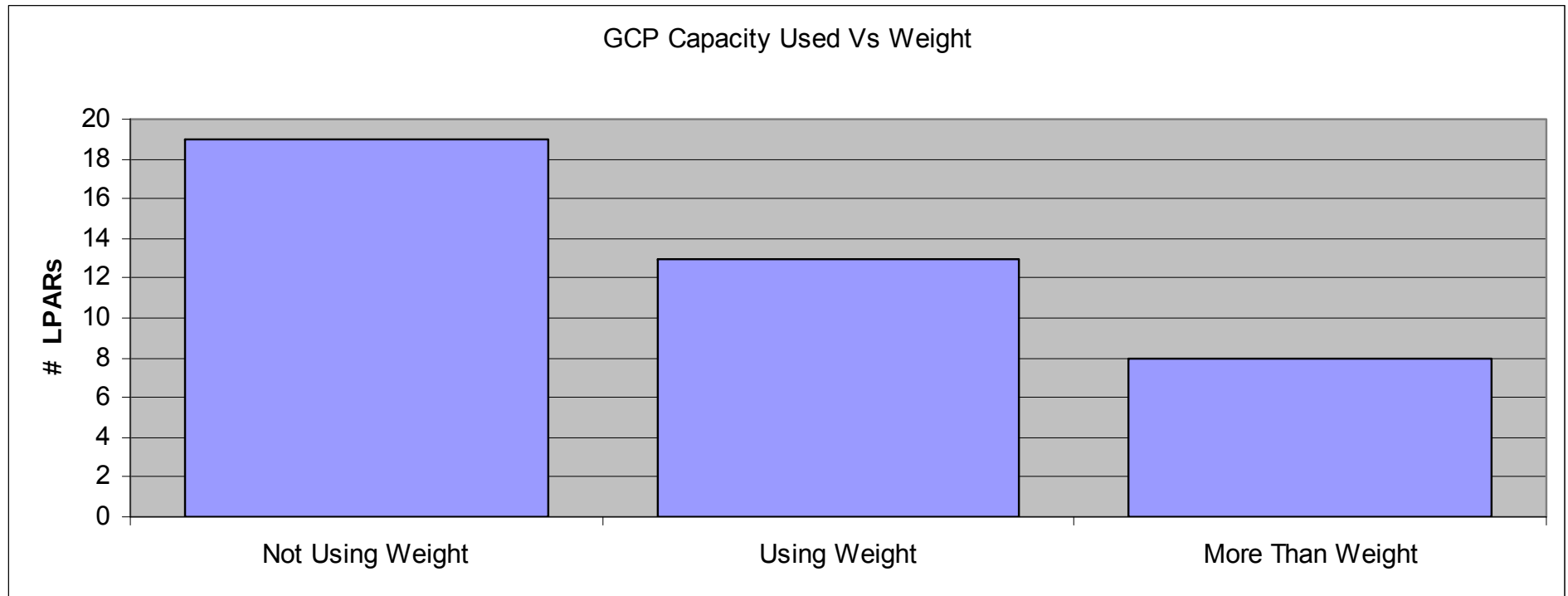
Total Processor GCP Busy



LPAR CPU Engines Distribution

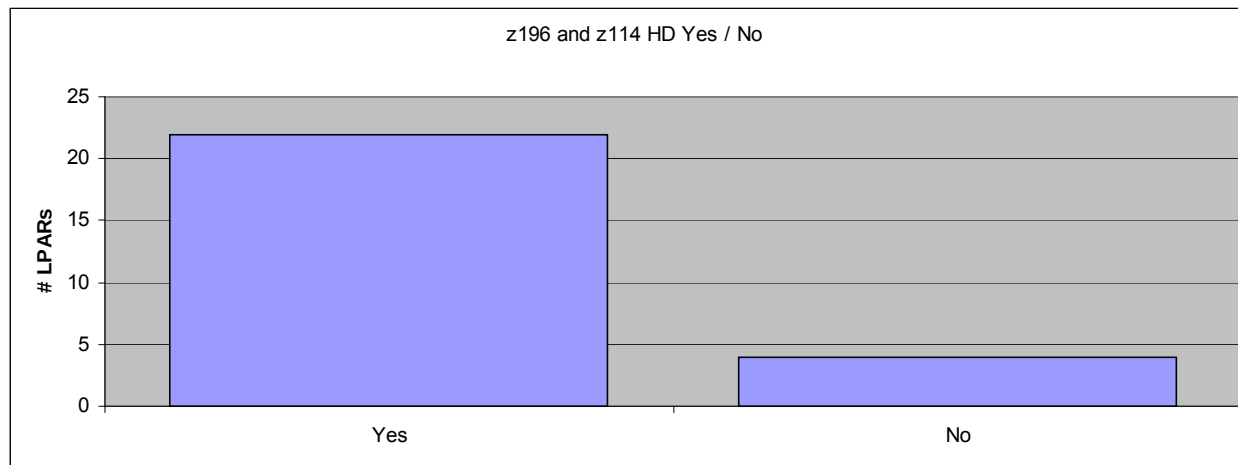
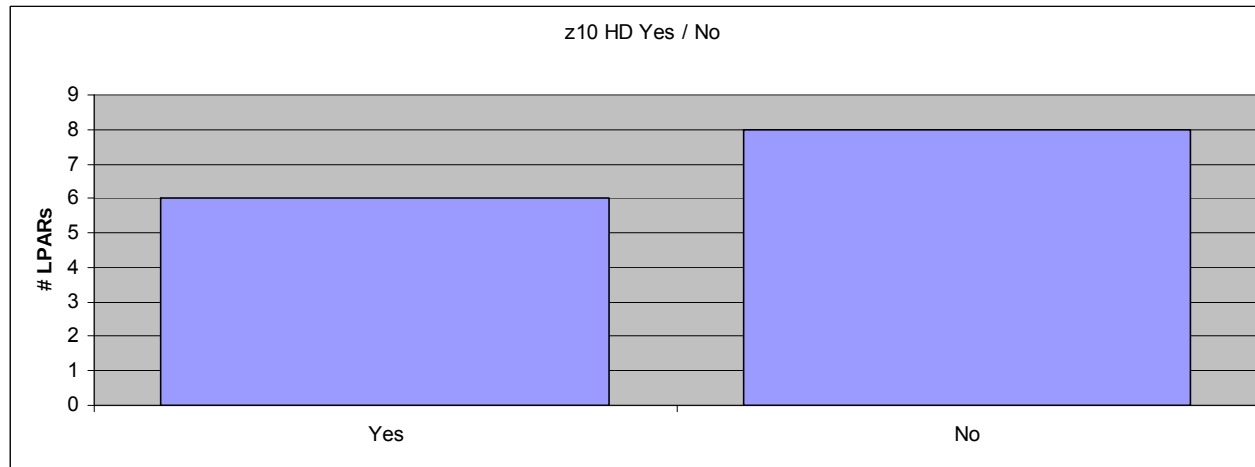


GCP Capacity Used Vs Weight



Not Using Weight	GCP Eng $\leq$ -.5 Weight
Using Weight	GCP Eng $>$ -.5 and $<$ +.5 Weight
More Than Weight	GCP Eng $\geq$ +.5 Weight

HD=Yes / No Distribution



- **z196**

- HD=YES is even more important on z196, ensure HD=YES, 0-11% for 1 Book z196
- See “Planning Considerations for HiperDispatch Mode **Version 2**” **WP101229**

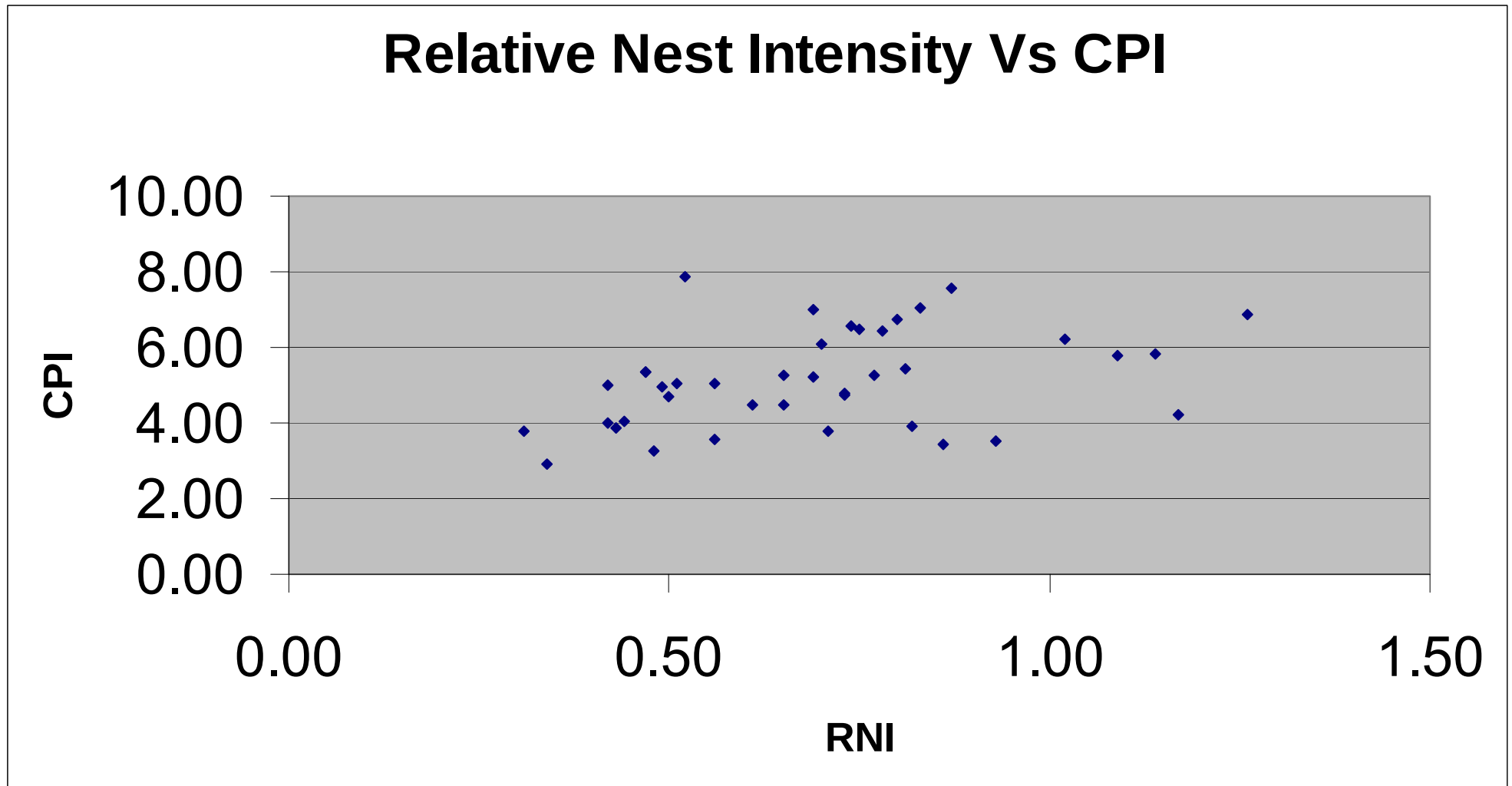
<http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101229>

CPU MF Averages – Technology Differences

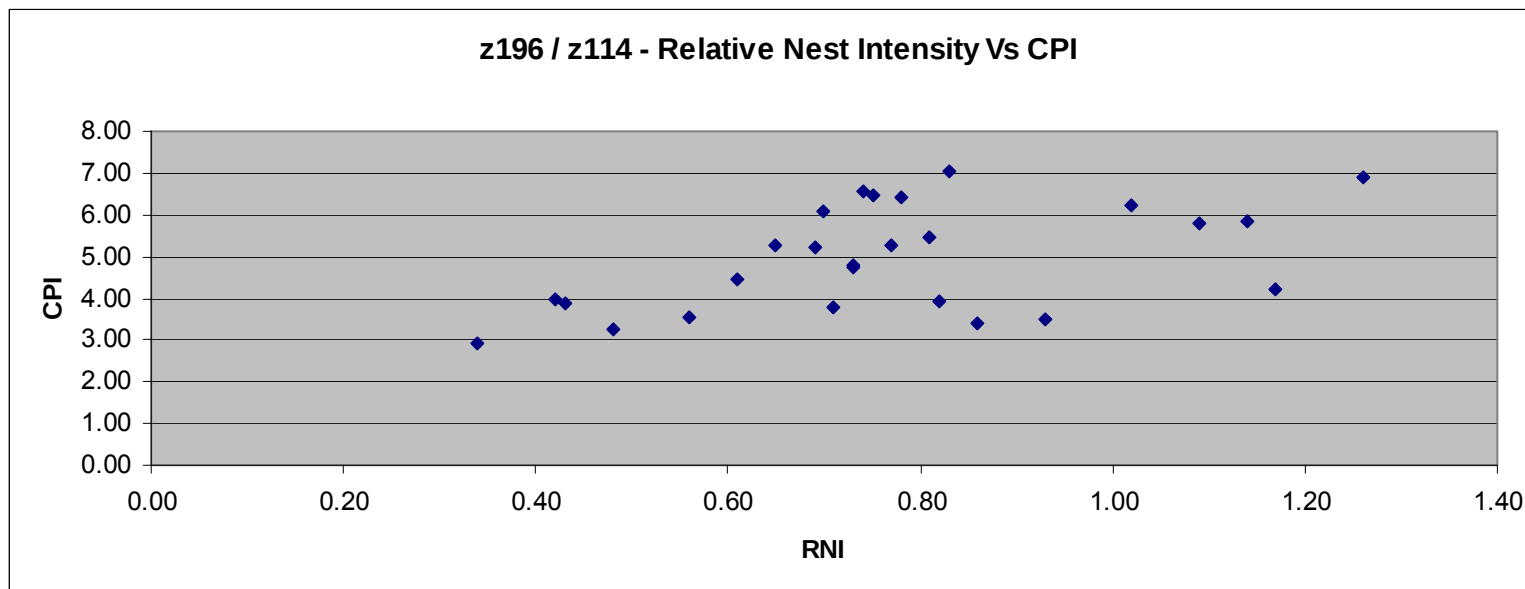
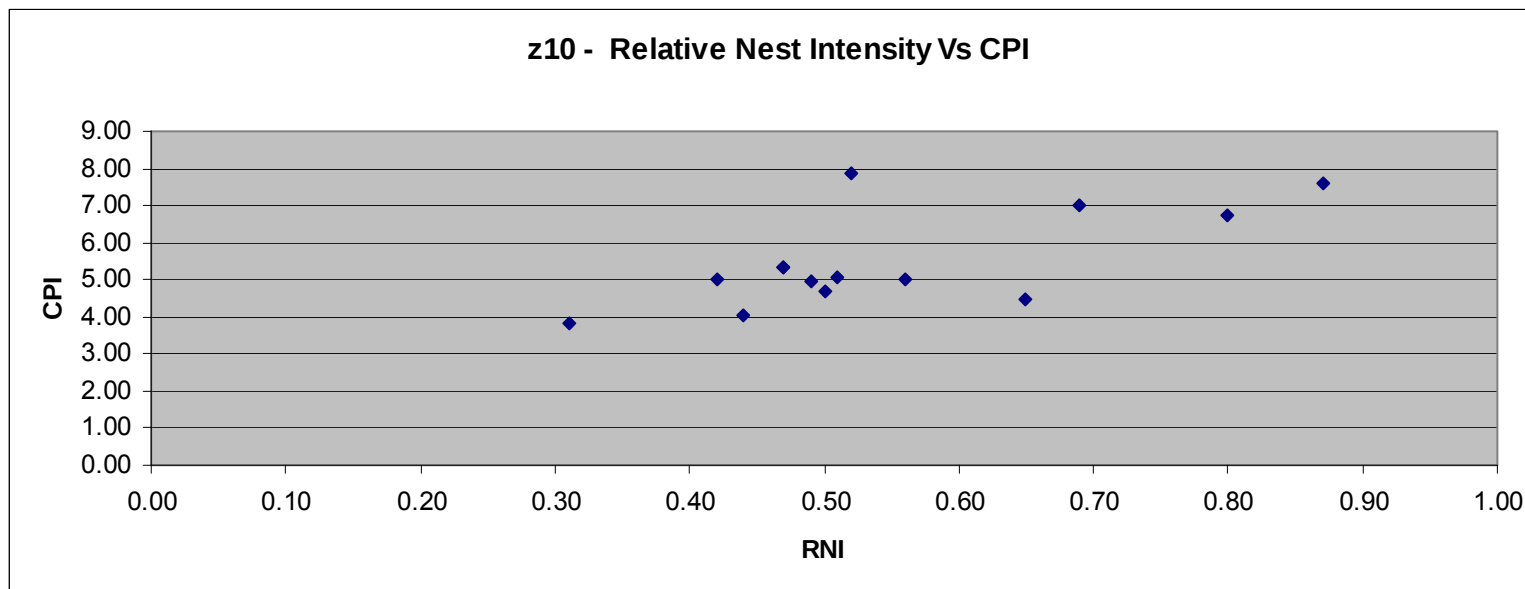
	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L15P / L2P	L3P	L2LP / L4LP	L2RP / L4RP	MEMP	Rel Nest Intensity	LPARCPU
z10	5.50	22.7	2.98	2.52	68	3.6	74.9	0.0	20.1	0.4	4.5	0.55	467.1
z196 / z114	4.96	35.5	2.55	2.41	63	3.8	60.8	23.8	11.3	0.7	3.4	0.77	425.4

CPU MF Metrics do not imply “goodness” or “badness”

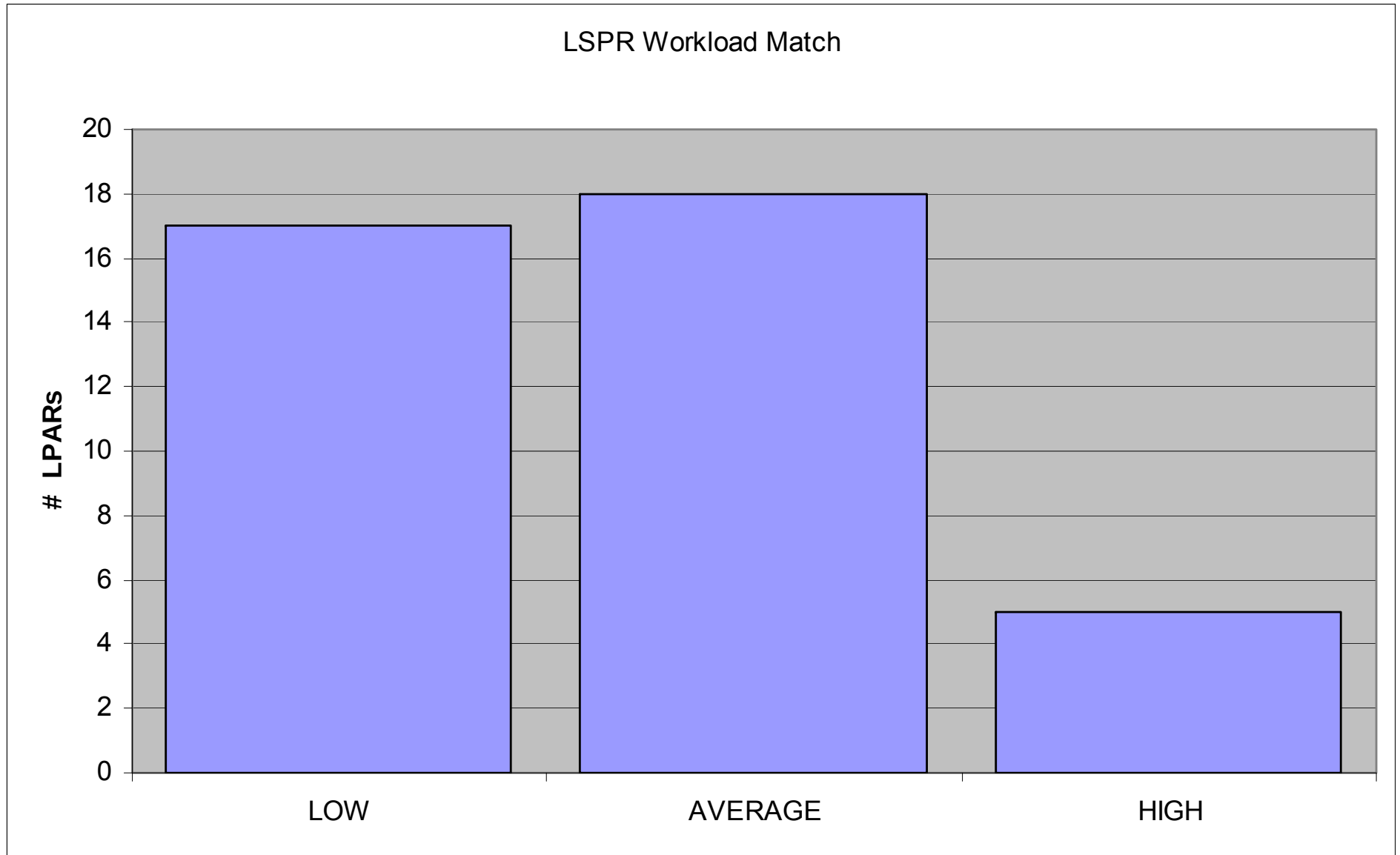
Data Profiles – RNI Vs CPI



Data Profiles – RNI Vs CPI



LSPR Workload Match



Data Profiles z10 and z196 / z114

			Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L15P / L2P	L3P	L2LP / L4LP	L2RP / L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Machine Type	LSPR Wkld	Machine	Model	HW Model	HD ?	CEC GCP Eng	LPAR Log GCP Eng	LPAR GCP Weight	Total Processor GCP %
SYSID	CPI	Prb State	CPI	CPI	SCPL1M	L1MP	L15P / L2P	L3P	L2LP / L4LP	L2RP / L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Machine Type	LSPR Wkld	Machine	Model	HW Model	HD ?	CEC GCP Eng	LPAR Log GCP Eng	LPAR GCP Weight	Total Processor GCP %
3	4.99	10.5	2.89	2.10	56	3.8	80.7	0.0	15.6	0.2	3.5	0.42	180.6	4.4	Z10	LOW	2097	604 E40	Yes	4	4	44.9%	72.6	
24	4.06	na	2.89	1.17	58	2.0	76.1	0.0	20.9	0.1	3.0	0.44	53.3	4.4	Z10	LOW	2097	706 E26	Yes	6	5	1.5%	82.3	
10	4.71	16.1	2.32	2.39	71	3.4	80.2	0.0	12.3	3.5	3.9	0.50	2335.6	4.4	Z10	LOW	2097	719 E40	Yes	19	17	94.0%	83.1	
7	5.05	19.1	2.71	2.34	63	3.7	77.2	0.0	18.4	0.0	4.4	0.51	280.1	4.4	Z10	LOW	2097	504 E12	Yes	4	4	66.0%	96.8	
1	7.86	33.0	4.03	3.82	68	5.7	69.6	0.0	27.0	0.1	3.3	0.52	362.1	3.2	Z10	LOW	2097	607 E26	Yes	7	6	100.0%	49.8	
36	7.02	21.2	2.99	4.03	89	4.5	71.3	0.0	20.7	2.3	5.7	0.69	575.6	4.4	Z10	AVG	2097	716 E40	Yes	16	8	39.0%	57.6	
30	3.80	na	2.71	1.08	48	2.3	82.0	0.0	16.0	0.0	1.9	0.31	442.7	4.4	Z10	LOW	2097	505 E12	No	5	5	DED	91.6	
4	5.33	24.2	3.01	2.32	58	4.0	76.1	0.0	20.4	0.0	3.5	0.47	473.2	4.4	Z10	LOW	2097	505 E12	No	5	5	91.0%	90.9	
29	5.36	na	3.01	2.36	58	4.0	77.3	0.0	19.0	0.0	3.7	0.47	272.4	4.4	Z10	LOW	2097	505 E12	No	5	4	86.0%	55.6	
8	4.97	na	2.87	2.10	61	3.5	80.0	0.0	15.5	0.0	4.5	0.49	125.4	2.1	Z10	LOW	2098	U02 E10	No	2	2	72.7%	60.5	
13	5.03	na	2.76	2.27	63	3.6	78.2	0.0	16.6	0.0	5.2	0.56	50.5	2.4	Z10	LOW	2098	V01 E10	No	1	1	73.1%	91.3	
12	4.49	6.1	2.75	1.74	69	2.5	66.3	0.0	28.9	0.0	4.9	0.65	351.0	4.4	Z10	LOW	2097	504 E12	No	4	4	70.0%	91.1	
23	6.76	na	3.20	3.56	94	3.8	67.2	0.0	25.4	0.1	7.3	0.80	884.7	4.4	Z10	AVG	2097	709 E26	No	9	8	100.0%	84.6	
18	7.58	51.3	3.52	4.06	95	4.3	66.7	0.0	25.1	0.0	8.2	0.87	151.7	4.4	Z10	AVG	2097	504 E12	No	4	3	23.5%	93.6	
5.50	22.7	2.98	2.52	68	3.6	74.9	0.0	20.1	0.4	4.5	0.55	467.1								7	5	66.3%	78.7	

			Est Instr Cmplx	Est Finite	Est		L15P /		L2LP /	L2RP /		Rel Nest			Machine	LSPR			HW		CEC	LPAR	LPAR	Total
SYSID	CPI	Prb State	CPI	CPI	SCPL1M	L1MP	L2P	L3P	L4LP	L4RP	MEMP	Intensity	LPARCPU	Eff GHz	Type	Wkld	Machine	Model	Model	HD ?	GCP Eng	Log GCP Eng	GCP Weight	Processor GCP %
34	3.98	40.7	2.36	1.62	46	3.6	72.1	19.5	6.3	0.7	1.4	0.42	1.8	3.8	Z114	LOW	2818	Z03 M10	Yes	3	3	14.3%	3.6	
40	3.55	na	2.07	1.48	50	2.9	68.8	19.8	7.9	1.4	2.1	0.56	1302.4	5.2	Z196	LOW	2817	716 M66	Yes	16	15	77.0%	76.1	
31	4.47	na	2.31	2.16	51	4.2	53.4	32.5	11.8	0.6	1.6	0.61	337.0	5.2	Z196	AVG	2817	704 M32	Yes	4	4	90.0%	77.0	
37	5.26	45.3	2.74	2.51	51	4.9	62.7	23.2	11.5	0.0	2.7	0.65	106.1	5.2	Z196	AVG	2817	504 M15	Yes	4	3	39.4%	51.7	
6	5.21	47.2	3.27	1.94	55	3.5	66.7	21.1	8.4	0.5	3.4	0.69	594.9	5.2	Z196	AVG	2817	734 M80	Yes	34	32	42.5%	33.4	
27	6.08	na	2.88	3.20	57	5.6	55.0	29.8	12.0	0.8	2.4	0.70	332.9	5.2	Z196	AVG	2817	704 M32	Yes	4	4	96.6%	76.5	
9	3.80	na	2.44	1.36	63	2.2	63.1	20.0	13.2	0.9	2.8	0.71	421.4	5.2	Z196	LOW	2817	510 M32	Yes	10	6	45.0%	73.1	
17	4.77	52.0	2.43	2.34	58	4.0	69.2	17.4	5.7	4.9	2.9	0.73	467.2	5.2	Z196	AVG	2817	716 M32	Yes	16	8	25.2%	69.6	
28	4.74	na	2.77	1.97	59	3.4	55.5	31.1	10.3	0.2	3.0	0.73	74.6	5.2	Z196	AVG	2817	604 M32	Yes	4	2	16.3%	92.1	
20	6.57	50.6	3.19	3.38	60	5.7	58.0	26.9	11.7	0.3	3.1	0.74	691.2	5.2	Z196	AVG	2817	709 M49	Yes	9	9	69.9%	96.2	
11	6.49	19.7	3.17	3.32	64	5.2	60.9	23.9	11.4	0.4	3.3	0.75	876.9	5.2	Z196	AVG	2817	720 M80	Yes	20	19	78.0%	49.9	
39	5.26	na	2.25	3.01	63	4.8	69.0	19.0	6.0	2.1	3.9	0.77	932.8	5.2	Z196	AVG	2817	724 M49	Yes	24	24	46.0%	78.8	
2	6.42	24.0	2.87	3.55	65	5.5	54.5	30.1	11.7	0.5	3.2	0.78	579.4	5.2	Z196	AVG	2817	610 M49	Yes	10	8	59.8%	63.7	
35	5.45	na	3.69	1.76	63	2.8	63.6	18.4	13.1	1.3	3.6	0.81	244.9	5.2	Z196	AVG	2817	707 M32	Yes	7	7	43.0%	99.9	
22	3.92	49.9	2.41	1.51	63	2.4	65.9	20.3	9.2	0.1	4.5	0.82	592.2	5.2	Z196	AVG	2817	711 M32	Yes	11	9	60.0%	68.3	
5	3.42	0.5	2.15	1.27	71	1.8	59.6	25.2	10.4	0.5	4.3	0.86	447.4	5.2	Z196	AVG	2817	710 M49	Yes	10	6	37.8%	68.2	
21	3.51	na	2.10	1.40	72	2.0	43.7	37.9	14.6	0.0	3.8	0.93	164.9	3.4	Z196	AVG	2817	607 M15	Yes	7	3	16.0%	73.8	
16	6.23	na	2.61	3.62	76	4.8	64.0	22.4	7.0	0.4	6.3	1.02	253.0	5.2	Z196	HIGH	2817	714 M66	Yes	14	6	25.6%	61.8	
33	5.80	19.9	1.81	3.98	83	4.8	41.8	31.5	22.1	0.3	4.4	1.09	216.5	5.2	Z196	HIGH	2817	714 M49	Yes	14	13	54.3%	71.8	
38	5.83	na	2.11	3.72	93	4.0	48.8	19.0	25.5	2.3	4.3	1.14	390.1	5.2	Z196	HIGH	2817	710 M32	Yes	10	7	32.5%	58.1	
14	4.23	na	2.25	1.98	91	2.2	52.7	26.2	14.5	0.2	6.4	1.17	49.5	5.2	Z196	AVG	2817	612 M32	Yes	12	3	5.6%	74.1	
15	6.88	40.8	2.71	4.16	95	4.4	52.6	27.9	11.9	0.2	7.4	1.26	412.7	5.2	Z196	HIGH	2817	711 M49	Yes	11	8	36.7%	78.6	
25	2.90	na	2.04	0.86	33	2.6	76.7	18.6	3.3	0.0	1.4	0.34	190.3	5.2	Z196	LOW	2817	704 M15	No	4	3	17.8%	82.9	
32	3.86	na	2.39	1.47	37	4.0	78.8	12.8	6.0	0.4	1.9	0.43	125.8	5.2	Z196	LOW	2817	707 M32	No	7	2	9.5%	74.6	
26	3.25	na	2.63	0.63	52	1.2	69.9	16.4	12.1	0.0	1.5	0.48	351.1	0.9	Z114	LOW	2818	M03 M05	No	3	3	93.2%	99.5	
19	7.04	36.2	2.66	4.38	69	6.4	53.5	27.8	15.0	0.4	3.3	0.83	903.2	5.2	Z196	HIGH	2817	511 M32	No	11	11	84.4%	94.8	
Avg	4.96	35.5	2.55	2.41	63	3.8	60.8	23.8	11.3	0.7	3.4	0.77	425.4							11	8	46.8%	71.1	

HiperDispatch Considerations

HiperDispatch Considerations

Adjusting Weight to increase Vertical Highs

- z196 710, Processor GCP Busy 58.1%, LPAR: 7 Logicals, 32.5% Weight
 - Results in 2 VHs, 2 VMs, and 3 VLs
 - Using more GCP than weight, 390.1 LPAR CPU Vs 325 weight (> .5 Engine more)
 - Data shows 2 VMs with lower L3Ps and higher RNIs than 2 VHs

- Consider assigning more weight from 32.5% to 35.1% to get additional VH
 - Processor Capacity is available
 - Fine tuning to increase L3P for Vertical High (thus lower RNI and Lower CPI)

CP ID Summary - SMF 113s

SMF 70s

CPID	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L15P / L2P	L3P	L2LP / L4LP	L2RP / L4RP	MEMP	Rel Nest Intensity	LPARCPU	PARKED	SHARE	%
0	6.14	0.0	2.09	4.04	95	4.3	47.1	20.4	25.5	2.4	4.6	1.19	87.7	0.0	100.0	HIGH
1	6.73	0.0	2.44	4.30	83	5.2	50.9	21.9	21.2	2.6	3.4	0.99	77.5	0.0	100.0	HIGH
2	6.26	0.0	2.03	4.23	101	4.2	46.4	17.4	29.3	2.4	4.6	1.22	80.8	0.0	62.5	MED
3	6.21	0.0	2.04	4.17	101	4.2	46.6	17.3	29.1	2.3	4.6	1.22	78.6	0.0	62.5	MED
4	5.59	0.0	2.34	3.25	90	3.6	49.5	19.3	25.0	2.0	4.1	1.10	21.6	74.4	0.0	LOW
5	5.62	0.0	2.35	3.26	88	3.7	49.0	19.7	25.5	1.9	3.9	1.08	17.4	79.3	0.0	LOW
6	5.50	0.0	2.37	3.13	84	3.7	50.6	19.7	24.1	1.8	3.7	1.03	14.3	83.0	0.0	LOW

HiperDispatch Considerations

- **z196 HiperDispatch=NO** specified for 4 LPARs
- **z196 Objective - keep VH Polarity Processors on same chip**
 - Source PU from On Chip L3 Cache
 - HD=YES is assumed LSPR / zPCR
- **L3 Off Chip and Off Book sourced from respective L4s**
 - CPU MF provides a measurement of this activity
- **Example from LPAR with HD=NO**
 - Opportunity cost: L4 Local sourcing that could have been resolved from L3

CPID	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L15P / L2P	L3P	<== L4LP that could have been L3P if HD=YES	L2LP / L4LP	L2RP / L4RP	MEMP	Rel Nest Intensity	LPARCPU
0	7.40	35.0	2.89	4.51	69	6.5	51.5	28.8	5.9	16.4	0.3	3.0	0.82	82.3
1	7.39	35.1	2.89	4.50	69	6.5	51.6	28.8	5.9	16.3	0.3	3.0	0.82	81.5
2	7.38	35.0	2.88	4.50	69	6.5	51.5	28.9	5.9	16.3	0.3	3.0	0.82	80.5
3	7.38	35.1	2.88	4.50	69	6.5	51.5	28.9	5.9	16.3	0.3	3.0	0.82	79.4
4	7.37	35.0	2.88	4.49	69	6.5	51.7	28.9	5.8	16.1	0.3	3.0	0.82	78.2
5	7.37	35.0	2.88	4.49	69	6.5	51.6	29.0	5.8	16.1	0.3	3.0	0.82	77.2
6	7.38	35.0	2.88	4.50	69	6.5	51.6	29.0	5.8	16.1	0.3	3.0	0.82	76.0
7	7.38	34.9	2.88	4.50	69	6.5	51.5	29.0	5.8	16.2	0.3	3.0	0.82	74.8
8	7.42	34.9	2.87	4.55	70	6.5	51.4	28.1	6.4	17.1	0.3	3.0	0.83	73.7
9	7.64	33.1	2.99	4.66	67	7.0	52.8	27.6	6.2	16.4	0.3	2.8	0.79	72.7
10	7.76	31.7	3.05	4.70	65	7.2	53.7	27.1	6.1	16.2	0.3	2.7	0.77	72.0

Summary

- **CPU MF Counters provide better information for more successful capacity planning**
- **Same data used to validate the LSPR workloads can now be obtained from production systems**
- **CPU MF Counters can also be useful for performance analysis**
- **Enable CPU MF Counters Today!**
 - Continuously collect SMF 113s for your production systems

References and Feedback

- **CPU MF Webinar Replays and Presentations**
 - <http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS4922>

- **Additional z/OS CPU MF information**
 - <http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TC000066>

- **How to Collect CPU Measurement Facility data for z/VM**
 - <http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105949>

Techdocs provides the latest ATS technical collateral

www.ibm.com/support/techdocs

The screenshot shows the IBM Techdocs website. At the top is the IBM logo and a navigation bar with links: Home, Solutions, Services, Products, Support & downloads, and My IBM. A search bar is located to the right of the navigation bar. Below the navigation bar, a welcome message reads "Welcome Kathy Walsh [Not you?] [IBM Sign in]". The main content area is titled "Techdocs - the Technical Sales Library". On the left side, there is a "Techdocs Library" sidebar with a list of categories: Flashes, Presentations & tools, Technotes & tips, FAQs, White papers, Solution scenario profiles, Customer support plans, Sizings, Auxiliary Material, Search Techdocs, and Techdocs feedback. Below this sidebar, there are "Related links" for Redbook publications and IBM Software Support Handbook. The main content area features a large image of a person working on a laptop. To the right of the image, there are three boxes: "New to Techdocs?" with a link to "Learn more", "Returning to Techdocs?" with a link to "Latest updates", and "Need Technical Support?" with a link to "Support & downloads". Below the image, a paragraph describes the site's purpose: "This site provides access to the Technical Sales Support organization's technical information databases. It gives you access to the most current installation, planning and technical support information available from IBM pre-sales support, and is constantly updated. You can browse or search these databases by date, document number, product, platform, keywords, etc." Below this paragraph, there is a section titled "New to Techdocs?" with a link to a "detailed introduction". Further down, there is a search section with a dropdown menu set to "All of the Techdocs Library", a search input field, and checkboxes for "Allow word variants" and "Fuzzy" search. The search results show "Hits: 50" and "Order by: relevance". There is also a link for "Help for Search". At the bottom, there is a section titled "Also available:" with a link to "Advanced search".

United States [change]

Search

Home Solutions Services Products Support & downloads My IBM

Welcome Kathy Walsh [Not you?] [IBM Sign in]

Techdocs - the Technical Sales Library

Techdocs Library

- Flashes
- Presentations & tools
- Technotes & tips
- FAQs
- White papers
- Solution scenario profiles
- Customer support plans
- Sizings
- Auxiliary Material
- Search Techdocs
- Techdocs feedback

Related links

- Redbook publications
- IBM Software Support Handbook

New to Techdocs? Take a look at our [detailed introduction](#), which describes the document categories available (those listed on the navigation area on the left side of this page).

Rather than browse these categories, as a convenience you may enter a search of the full **Techdocs** database, or of any category you wish, here:

Search: All of the Techdocs Library ☐ Allow word variants

for: ☐ "Fuzzy" search

Hits: 50 Order by: relevance

Include docs updated: any time [Help for Search](#)

Also available: our [Advanced search](#), where you can select documents based on various assigned document attributes.

New to Techdocs?

Is this your first visit to **Techdocs** (the Technical Sales Library)?

→ [Learn more](#)

Returning to Techdocs?

Looking for what's new in the **Techdocs Library**?

→ [Latest updates](#)

Need Technical Support?

Looking for support resources or other documents and tools?

→ [Support & downloads](#)

System z Social Media

- System z official Twitter handle:

- ▶ [@ibm_system_z](https://twitter.com/ibm_system_z)

- Top Facebook pages related to System z:

- ▶ [Systemz Mainframe](#)
 - ▶ [IBM System z on Campus](#)
 - ▶ [IBM Mainframe Professionals](#)
 - ▶ [Millennial Mainframer](#)

- Top LinkedIn Groups related to System z:

- ▶ [Mainframe Experts Network](#)
 - ▶ [Mainframe](#)
 - ▶ [IBM Mainframe](#)
 - ▶ [System z Advocates](#)
 - ▶ [Cloud Mainframe Computing](#)

- YouTube

- ▶ [IBM System z](#)



- Leading Blogs related to System z:

- ▶ [Evangelizing Mainframe \(Destination z blog\)](#)
 - ▶ [Mainframe Performance Topics](#)
 - ▶ [Common Sense](#)
 - ▶ [Enterprise Class Innovation: System z perspectives](#)
 - ▶ [Mainframe](#)
 - ▶ [MainframeZone](#)
 - ▶ [Smarter Computing Blog](#)
 - ▶ [Millennial Mainframer](#)

Thank You for Attending!



Trademarks

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

AlphaBlox*	GDPS*	RACF*	Tivoli*
APPN*	HiperSockets	Redbooks*	Tivoli Storage Manager
CICS*	HyperSwap	Resource Link	TotalStorage*
CICS/VSE*	IBM*	RETAIN*	VSE/ESA
Cool Blue	IBM eServer	REXX	VTAM*
DB2*	IBM logo*	RMF	WebSphere*
DFSMS	IMS	S/390*	zEnterprise
DFSMSHsm	Language Environment*	Scalable Architecture for Financial Reporting	xSeries*
DFSMSrmm	Lotus*	Sysplex Timer*	z9*
DirMaint	Large System Performance Reference™ (LSPR™)	Systems Director Active Energy Manager	z10
DRDA*	Multiprise*	System/370	z10 BC
DS6000	MVS	System p*	z10 EC
DS8000	OMEGAMON*	System Storage	z/Architecture*
ECKD	Parallel Sysplex*	System x*	z/OS*
ESCON*	Performance Toolkit for VM	System z	z/VM*
FICON*	PowerPC*	System z9*	z/VSE
FlashCopy*	PR/SM	System z10	zSeries*
	Processor Resource/Systems Manager		

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* 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.

Back Up

Formulas – z10

Workload Characterization
L1 Sourcing from cache/memory hierarchy

Metric	Calculation – <i>note all fields are deltas between intervals</i>
CPI	$B0 / B1$
PRBSTATE	$(P33 / B1) * 100$
L1MP	$((B2+B4) / B1) * 100$
L15P	$((E128+E129) / (B2+B4)) * 100$
L2LP	$((E130+E131) / (B2+B4)) * 100$
L2RP	$((E132+E133) / (B2+B4)) * 100$
MEMP	$((E134+E135) + (B2+B4-E128-E129-E130-E131-E132-E133-E134-E135)) / (B2+B4)) * 100$
LPARCPU	$(((1/CPSP/1,000,000) * B0) / \text{Interval in Seconds}) * 100$

CPI – Cycles per Instruction

PRBSTATE - % Problem State

L1MP – Level 1 Miss Per 100 instructions

L15P – % sourced from L1.5 cache

L2LP – % sourced from Level 2 Local cache (on same book)

L2RP – % sourced from Level 2 Remote cache (on different book)

MEMP - % sourced from Memory

LPARCPU - APPL% (GCPs, zAAPs, zIIPs) captured and uncaptured

B* - Basic Counter Set - Counter Number

P* - Problem-State Counter Set - Counter Number

See “The Set-Program-Parameter and CPU-Measurement Facilities”
SA23-2260-0 for full description

E* - Extended Counters - Counter Number

See “IBM The CPU-Measurement Facility Extended Counters Definition
for z10” SA23-2261-0/1 for full description

CPSP - SMF113_2_CPSP “CPU Speed”

Formulas – z10 Additional

Metric	Calculation – <i>note all fields are deltas between intervals</i>
Est Instr Cmplx CPI	CPI – Estimated Finite CPI
Est Finite CPI	$((B3+B5) / B1) * .84$
Est SCPL1M	$((B3+B5) / (B2+B4)) * .84$
Rel Nest Intensity	$(1.0*L2LP + 2.4*L2RP + 7.5*MEMP) / 100$
Eff GHz	CPSP / 1000

Note these Formulas may change in the future

Est Instr Cmplx CPI – Estimated Instruction Complexity CPI (infinite L1)

Est Finite CPI – Estimated CPI from Finite cache/memory

Est SCPL1M – Estimated Sourcing Cycles per Level 1 Miss

Rel Nest Intensity – Reflects distribution and latency of sourcing from shared caches and memory

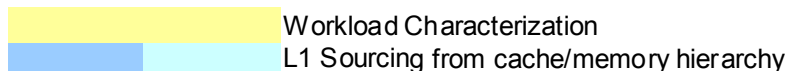
Eff GHz – Effective gigahertz for GCPs, cycles per nanosecond

B* - Basic Counter Set - Counter Number

P* - Problem-State Counter Set - Counter Number

See “The Set-Program-Parameter and CPU-Measurement Facilities”
SA23-2260-0 for full description

CPSP - SMF113_2_CPSP “CPU Speed”



Formulas – z196

Workload Characterization

L1 Sourcing from cache/memory hierarchy

Metric	Calculation – <i>note all fields are deltas between intervals</i>
CPI	$B0 / B1$
PRBSTATE	$(P33 / B1) * 100$
L1MP	$((B2+B4) / B1) * 100$
L2P	$((E128+E129) / (B2+B4)) * 100$
L3P	$((E150+E153) / (B2+B4)) * 100$
L4LP	$((E135+E136+E152+E155) / (B2+B4)) * 100$ L3 off chip in local L4
L4RP	$((E138+E139+E134+E143) / (B2+B4)) * 100$ L3 off book in remote L4
MEMP	$((E141+E142) + (B2+B4-E128-E129-E150-E153-E135-E136-E152-E155-E138-E139-E134-E143-E141-E142)) / (B2+B4)) * 100$
LPARCPU	$((1/CPSP/1,000,000) * B0) / \text{Interval in Seconds} * 100$

CPI – Cycles per Instruction

Prb State - % Problem State

L1MP – Level 1 Miss Per 100 instructions

L2P – % sourced from Level 2 cache

L3P – % sourced from Level 3 on same Chip cache

L4LP – % sourced from Level 4 Local cache (on same book)

L4RP – % sourced from Level 4 Remote cache (on different book)

MEMP - % sourced from Memory

LPARCPU - APPL% (GCPs, zAAPs, zIIPs) captured and uncaptured

B* - Basic Counter Set - Counter Number

P* - Problem-State Counter Set - Counter Number

See “The Set-Program-Parameter and CPU-Measurement Facilities”
SA23-2260-0 for full description

E* - Extended Counters - Counter Number

See “The CPU-Measurement Facility Extended Counters Definition for
z10 and z196” SA23-2261-01 for full description

CPSP - SMF113_2_CPSP “CPU Speed”

Formulas – z196 Additional

Metric	Calculation – <i>note all fields are deltas between intervals</i>
Est Instr Cmplx CPI	CPI – Estimated Finite CPI
Est Finite CPI	$((B3+B5) / B1) * (.59 + (0.1*RNI))$ updated *
Est SCPL1M	$((B3+B5) / (B2+B4)) * (.59 + (0.1*RNI))$ updated *
Rel Nest Intensity	$1.67 * (0.4*L3P + 1.0*L4LP + 2.4*L4RP + 7.5*MEMP) / 100$ updated *
Eff GHz	CPSP / 1000

Note these Formulas may change in the future

* Updated July 2012

Est Instr Cmplx CPI – Estimated Instruction Complexity CPI (infinite L1)

Est Finite CPI – Estimated CPI from Finite cache/memory

Est SCPL1M – Estimated Sourcing Cycles per Level 1 Miss

Rel Nest Intensity – Reflects distribution and latency of sourcing from shared caches and memory

Eff GHz – Effective gigahertz for GCPs, cycles per nanosecond

B* - Basic Counter Set - Counter Number

P* - Problem-State Counter Set - Counter Number

See “The Set-Program-Parameter and CPU-Measurement Facilities”
SA23-2260-0 for full description

CPSP - SMF113_2_CPSP “CPU Speed”

Workload Characterization
L1 Sourcing from cache/memory hierarchy

Formulas – Additional TLB

Metric – z10	Calculation – <i>note all fields are deltas between intervals</i>
Est. TLB1 CPU Miss % of Total CPU	$((E145+E146) / B0) * 100 * .31 *$
Estimated TLB1 Cycles per TLB Miss	$(E145+E146) / (E138+E139) * .31 *$
PTE % of all TLB1 Misses	$(E140 / (E138+E139)) * 100$

Metric – z196	Calculation – <i>note all fields are deltas between intervals</i>
Est. TLB1 CPU Miss % of Total CPU	$((E130+E131) / B0) * 100 * .47 *$
Estimated TLB1 Cycles per TLB Miss	$(E130+E131) / (E144+E145) * .47 *$
PTE % of all TLB1 Misses	$(E146 / (E144+E145)) * 100$

Note these Formulas may change in the future

* Updated March 2012

Est. TLB1 CPU Miss % of Total CPU - Estimated TLB CPU % of Total CPU B* - Basic Counter Set - Counter Number

Estimated TLB1 Cycles per TLB Miss – Estimated Cycles per TLB Miss

PTE % of all TLB1 Misses – Page Table Entry % misses

See “The Set-Program-Parameter and CPU-Measurement Facilities” SA23-2260-0 for full description

E* - Extended Counters - Counter Number

See “IBM The CPU-Measurement Facility Extended Counters Definition for z10” SA23-2261-0 for full description or “The CPU-Measurement Facility Extended Counters Definition for z10 and z196” SA23-2261-01 for full description