2015 CPU MF Update

John Burg
IBM

March 3, 2015
Session Number 16803

#SHAREorg

SHARE is an independent volunteer-run information technology association that provides education, professional networking and industry influence.

SHARE in Seattle 2015

Trademarks

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

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a more complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*BladeCenter®, CICS®, DataPower®, DB2®, e business(logo)®, ESCON, eServer, FICON®, IBM®, IBM (logo)®, IMS, MVS, OS/390®, POWER6®, POWER6+, POWER7®, Power Architecture®, PowerVM®, PureFlex, PureSystems, S/390®, ServerProven®, Sysplex Timer®, System p®, System p5, System x®, z Systems®, System z9®, System z10®, WebSphere®, X-Architecture®, z13™, z Systems™, z9®, z10, z/Architecture®, z/OS®, z/VM®, z/VSE®, zEnterprise®, zSeries®

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 Sync 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 Sync 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.

Notice Regarding Specialty Engines (e.g., zIIPs, zAAPs and IFLs):

Any information contained in this document regarding Specialty Engines ("SEs") and SE eligible workloads provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs). IBM authorizes customers to use IBM SEs only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the “Authorized Use Table for IBM Machines” provided at:
www.ibm.com/systems/support/machine_warranties/machine_code/aut.html (“AUT”).

No other workload processing is authorized for execution on an SE.

IBM offers SEs at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

Agenda – CPU MF Counters

- Value of CPU MF Counters
  - What, Why important for z13, and How to implement
- What's New
  - zEC12 RNI
  - z13
    - Display
    - Metrics and Formulas
    - Topology - SMF 99 Subtype 14s
    - WSC Tests

      SMT-2 and SMT-1 Examples
- Looking for z13 Volunteers
- Summary

Value of CPU MF Measurement Facility (CPU MF)

- Recommended Methodology for successful z Systems Processor Capacity Planning
  - Need on “Before” processor to determine LSPR workload
- Validate achieved z Systems processor performance
  - Needed on “Before” and “After” processors
- Provide insights for new features and functions
  - Continuously running on all LPARs

Capturing CPU MF data is an Industry “Best Practice”

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
- 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

z Systems Capacity Planning

Processor	Features	Flag	MSU	LSPR Workload Category				
				Low	Low-Avg	Average	Avg-High	High
z13/700								
2964-701	1W	=	210	1,779	1,736	1,695	1,614	1,540
2964-702	2W	=	394	3,452	3,319	3,196	3,003	2,833
2964-703	3W	=	571	5,085	4,854	4,644	4,340	4,073
2964-704	4W	=	740	6,678	6,344	6,041	5,625	5,262
2964-705	5W	=	905	8,238	7,792	7,392	6,866	6,410
2964-706	6W	=	1,062	9,765	9,202	8,700	8,066	7,518
2964-707	7W	=	1,212	11,260	10,573	9,964	9,224	8,587
2964-708	8W	=	1,356	12,724	11,906	11,188	10,344	9,618
2964-709	9W	=	1,496	14,157	13,204	12,371	11,425	10,613
2964-710	10W	=	1,632	15,560	14,466	13,515	12,469	11,574

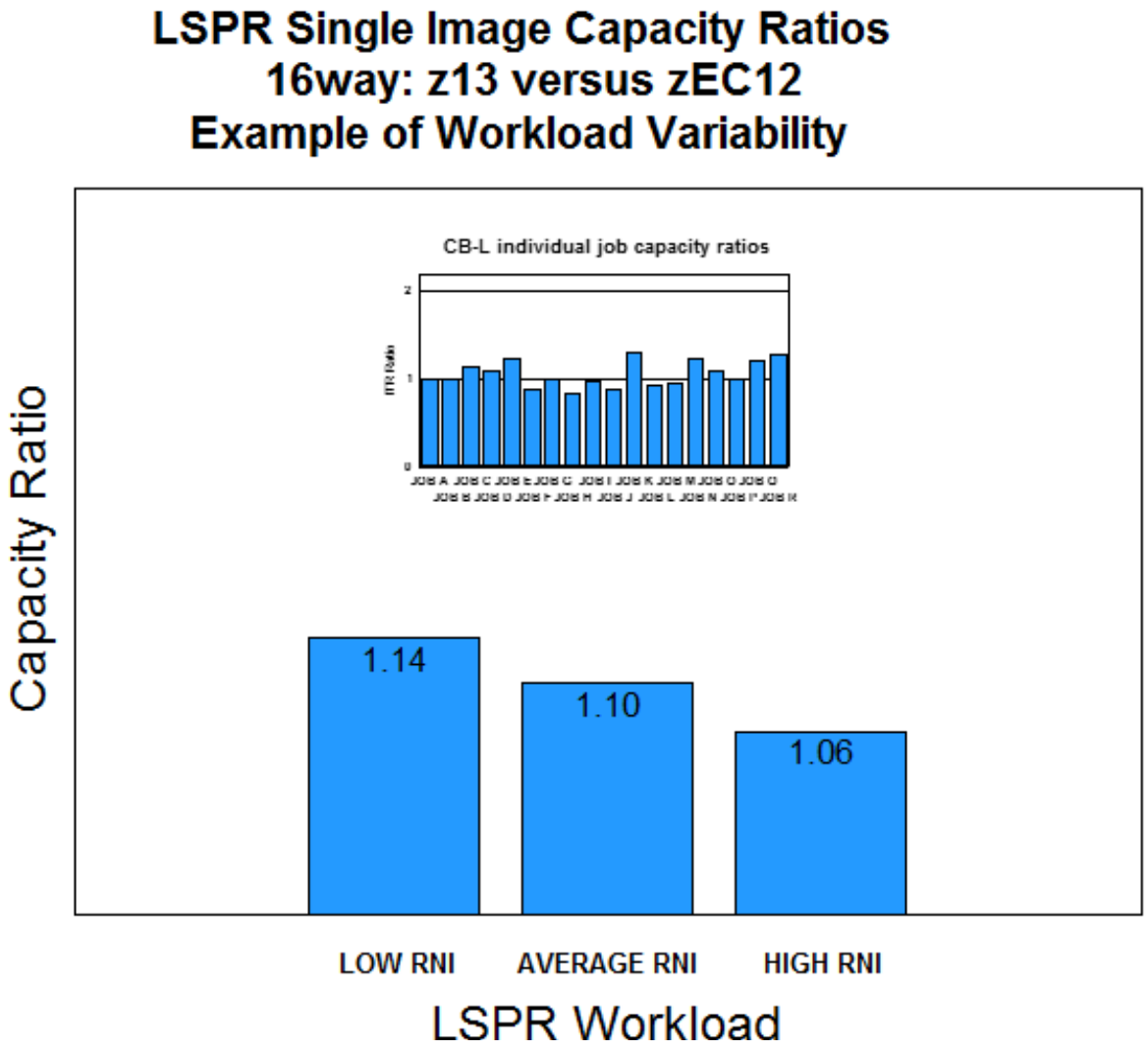
- Relative Processor Capacity varies by LPAR configuration and Workload
- CPU MF data used to select LSPR Workload Match
- IBM Capacity Planning Tools utilize CPU MF data to select a workload
  - zPCR, CP3000 and zBNA are all enabled for CPU MF

z13 Processor Performance

- New Processor Design
  - Includes major pipeline enhancements and Larger Caches
  - 1.10x (10%) average performance improvement at equal Nway Vs zEC12

- Workload Variability
  - Workloads moving to z13 may see more variability than the last few migrations
  - Potential Sources of Variability
    - Workload interaction with Processor Design – may have variable but unpredictable benefit to IPC
    - PR/SM placement of CPs and memory for an LPAR

LSPR Single Image Capacity Ratios - 16 Way z13 versus zEC12



Additional Customer Value with CPU MF Counters data

- 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
  - Supported by many software products including Tivoli TDSz
- Some examples of usage include:
  - Impact zEDC compression
  - HiperDispatch Impact
  - Configuration changes (Additional LPARs)
  - 1 MB Page implementation
  - Application Changes (e.g. CICS Threadsafe Vs QR)
  - Estimating Utilization Effect for capacity planning
  - GHz change in Power Saving Mode
  - Crypto CPACF usage

CPU MF Counters Enablement Resources

- CPU MF Webinar Replays and Presentations
  - http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS4922
- z/OS CPU MF - “Detailed Instructions” Step by Step Guide
  - http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TC000066
- z/VM Using CPU Measurement Facility Host Counters
  - http://www.vm.ibm.com/perf/tips/cpumf.html

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**

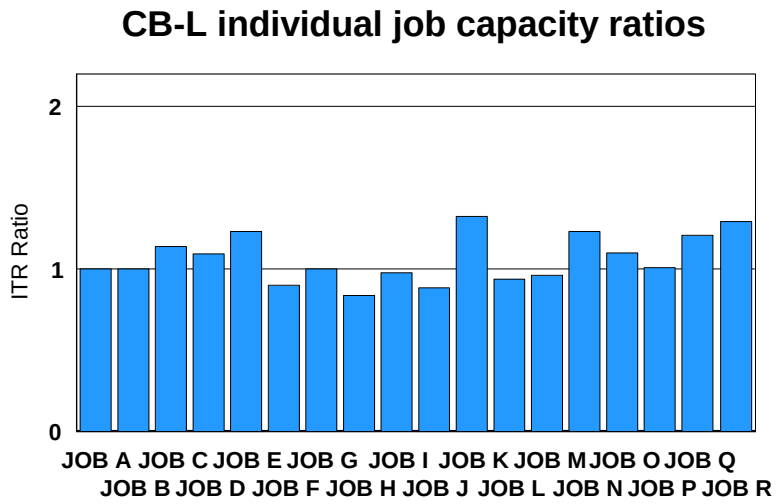
SMF 113s Space Requirements Are Minimal

- 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%

Processor Measurement Methodology

- zPCR provides overall processor capacity expectations
- Variations within workload is expected
  - An individual job can see a shortfall but the measurement is for the entire workload
- Take care with customer synthetic benchmarks as they are often subject to significant measurement error
- See "*Processor Migration Capacity Analysis in a Production Environment* "
  - http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP100744



CPU MF Update

zEC12 and z13

CPU MF – What's New for z13?

- Updated zEC12 RNI
- Same LSPR RNI Workload Decision Table
- Same Metrics as previous processors
  - New formulas
- New "Miss" cycles measurement allows improved Metrics:
  - CPI = Instruction Complexity CPI + Finite Cache CPI
  - Estimated Sourcing Cycles per L1 Miss
- SMF 99s Subtype 14
  - Drawer, Node, Chip identification for Logical Processor (thread)
- CPU MF Metrics at Logical Processor or Thread level
  - SMT 1 Logical Processor level
  - SMT 2 Thread level
- New MT Diagnostic Counter Set

Formulas – zEC12 / zBC12 Additional

Metric	**Calculation** – *note all fields are **deltas** between intervals*
Est Instr Cmplx CPI	CPI – Estimated Finite CPI
Est Finite CPI	((B3+B5) / B1) * (.54 + (0.04*RNI))
Est SCPL1M	((B3+B5) / (B2+B4)) * (.54 + (0.04*RNI))
Rel Nest Intensity	2.3*(0.4*L3P + 1.2*L4LP + 2.7*L4RP + 8.2*MEMP) / 100
Eff GHz	CPSP / 1000

Updated January 2015

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

Workload Characterization

L1 Sourcing from cache/memory hierarchy

B* - Basic Counter Set - Counter Number

P* - Problem-State Counter Set - Counter Number

See “The Load-Program-Parameter and CPU-Measurement Facilities” SA23-2260-03 for full description

CPSP - SMF113_2_CPSP “CPU Speed”

RNI-based LSPR Workload Decision Table

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

Current table applies to z10 EC, z10 BC, z196, z114, zEC12, zBC12 and z13 CPU MF data

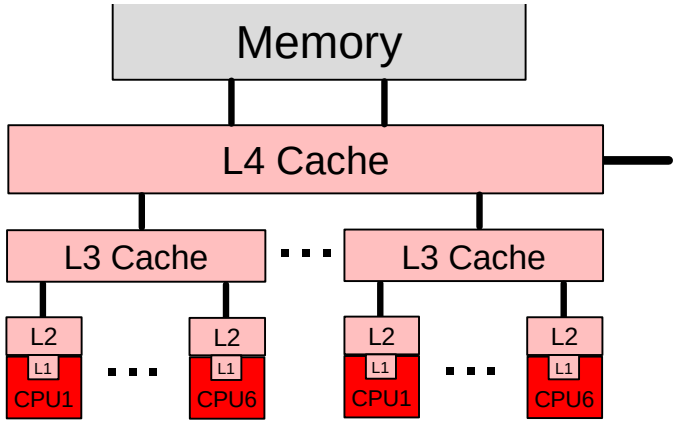
z13 Metrics

IBM z13 versus zEC12 Hardware Comparison

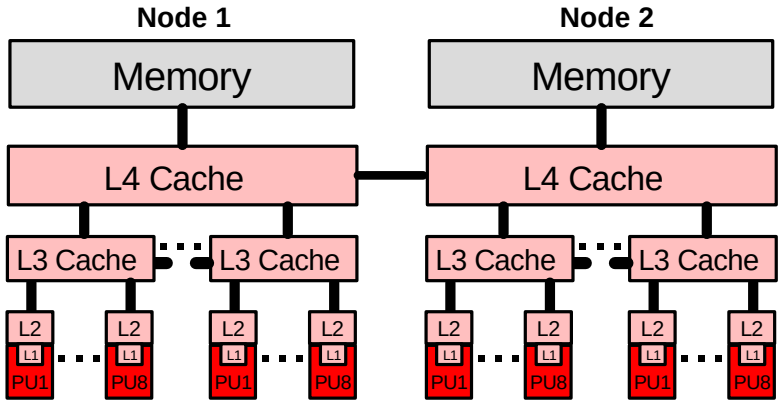
- zEC12
  - CPU
    - 5.5 GHz
    - Enhanced Out-Of-Order
  - Caches
    - L1 private 64k i, 96k d
    - L2 private 1 MB i + 1 MB d
    - L3 shared 48 MB / chip
    - L4 shared 384 MB / book
- z13
  - CPU
    - 5.0 GHz
    - Major pipeline enhancements
  - Caches
    - L1 private 96k i, 128k d
    - L2 private 2 MB i + 2 MB d
    - L3 shared 64 MB / chip
    - L4 shared 480 MB / node

      - plus 224 MB L3 NIC Directory

Single Book View



Single Drawer View - Two Nodes



z/OS SMF 113 Record

- SMF113_2_CTRVN2
  - "1" = z10
  - "2" = z196 / z114
  - "3" = zEC12 / zBC12
  - "4" = z13

Operations – Display HIS Command on z13

```
COMMAND INPUT ===> _                                          SCROLL ===
RESPONSE=SYSD
 HIS015I 17.48.11 DISPLAY HIS 822
 HIS      002A ACTIVE
 COMMAND: MODIFY HIS,B,TT='CMU MF COUNTERS ENABLED',CTRONLY,CTR=ALL,SI=
          SYNC
 START TIME:  2015/02/24 15:34:08
 END TIME:    ----/--/-- --:--:--
 COMPLETION STATUS: --------
 FILE PREFIX: SYSHIS20150224.153408.
 COUNTER VERSION NUMBER 1: 1   COUNTER VERSION NUMBER 2: 4
 COMMAND PARAMETER VALUES USED:
  TITLE=  CMU MF COUNTERS ENABLED
  PATH=   .
  COUNTER SET= BASIC,PROBLEM-STATE,CRYPTO-ACTIVITY,EXTENDED,
               MT-DIAGNOSTIC
  DURATION= NOLIMIT
  CTRONLY
  DATALOSS= IGNORE
```

z13 “4”

Formulas – z13

Workload Characterization
L1 Sourcing from cache/memory hierarchy

Metric	Calculation – *note all fields are deltas. SMF113-1s are deltas. SMF 113-2s are cumulative.*
CPI	B0 / B1
PRBSTATE	(P33 / B1) * 100
L1MP	((B2+B4) / B1) * 100
L2P	((E133+E136) / (B2+B4)) * 100
L3P	((E144+E145+ E162+E163) / (B2+B4)) * 100
L4LP	((E146+E147+E148+E164+E165+E166) / (B2+B4)) * 100
L4RP	((E149+E150+E151+E152+E153+E154+E155+E156+E157+E167+ E168+E169+E170+E171+E172+ E173+E174+E175) / (B2+B4)) * 100
MEMP	((E158 + E159 + E160 + E161 + E176 + E177 + E178 + E179) / (B2+B4)) * 100
LPARCPU	(((1/CPSP/1,000,000) * B0) / 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 Load-Program-Parameter and CPU-Measurement Facilities” SA23-2260 for full description

E* - Extended Counters - Counter Number

See “IBM The CPU-Measurement Facility Extended Counters Definition for z10, z196/ z114, zEC12 /zBC12 and z13” SA23-2261-03 for full description

CPSP - SMF113_2_CPSP “CPU Speed”

Note these Formulas may change in the future

Formulas – z13 Additional

Metric	**Calculation**– *note all fields are deltas. SMF113-1s are deltas. SMF 113-2s are cumulative.*
Est Instr Cmplx CPI	CPI – Estimated Finite CPI
Est Finite CPI	E143 / B1
Est SCPL1M	E143 / (B2+B4)
Rel Nest Intensity	2.6*(0.4*L3P + 1.6*L4LP + 3.5*L4RP + 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

Workload Characterization
L1 Sourcing from cache/memory hierarchy

B* - Basic Counter Set - Counter Number

P* - Problem-State Counter Set - Counter Number

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

E* - Extended Counters - Counter Number

See “IBM The CPU-Measurement Facility Extended Counters Definition for z10, z196/ z114, zEC12 /zBC12 and z13” SA23-2261-03 for full description

CPSP - SMF113_2_CPSP “CPU Speed”

Formulas – z13 Additional TLB

Metric	**Calculation** – *note all fields are deltas. SMF113-1s are deltas. SMF 113-2s are cumulative.*
Est. TLB1 CPU Miss % of Total CPU	((E130+E135) / B0) * (E143 / (B3+B5)) *100
Estimated TLB1 Cycles per TLB Miss	(E130+E135) / (E129+E134) * (E143 / (B3+B5))
PTE % of all TLB1 Misses	(E137 / (E129+E134)) * 100
TLB Miss Rate	(E129 + E134) / interval

Note these Formulas may change in the future

Est. TLB1 CPU Miss % of Total CPU - Estimated TLB CPU % of Total CPU

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

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

TLB Miss Rate – TLB Misses per interval (interval is defined by user for length of measurement and units)

B* - Basic Counter Set - Counter Number

P* - Problem-State Counter Set - Counter Number

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

E* - Extended Counters - Counter Number

See “IBM The CPU-Measurement Facility Extended Counters Definition for z10, z196/ z114, zEC12 /zBC12 and z13” SA23-2261-03 for full description

CPSP - SMF113_2_CPSP “CPU Speed”

Sample WSC z13 CPU MF Metrics

CPU MF - zIIPs

Hour	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L2P	L3P	L4LP	L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Machine Type	LSPR Wkld
10:10	1.92	95.9	1.47	0.45	40	1.1	74.6	13.8	4.7	2.8	4.1	1.40	478.6	5.0	Z13	AVG
10:15	1.93	95.4	1.49	0.44	40	1.1	73.8	13.6	5.2	2.9	4.5	1.50	429.8	5.0	Z13	AVG
10:25	1.63	95.4	1.15	0.48	52	0.9	67.6	16.5	6.7	3.7	5.6	1.87	359.7	5.0	Z13	AVG
10:30	1.64	95.4	1.15	0.49	52	0.9	67.4	16.7	6.7	3.7	5.5	1.86	361.2	5.0	Z13	AVG
10:40	1.93	95.4	1.49	0.44	41	1.1	73.9	13.5	5.2	2.9	4.6	1.51	427.4	5.0	Z13	AVG
10:45	1.93	95.3	1.49	0.44	41	1.1	73.7	13.6	5.3	2.9	4.5	1.51	427.4	5.0	Z13	AVG
10:55	1.62	95.3	1.17	0.45	48	0.9	69.6	14.8	6.9	3.3	5.4	1.79	326.8	5.0	Z13	AVG
11:00	1.61	95.4	1.17	0.45	48	0.9	69.8	14.7	6.9	3.3	5.4	1.78	325.4	5.0	Z13	AVG

Workload Characterization
L1 Sourcing from cache/memory hierarchy

SMF 99s (subtype 14)

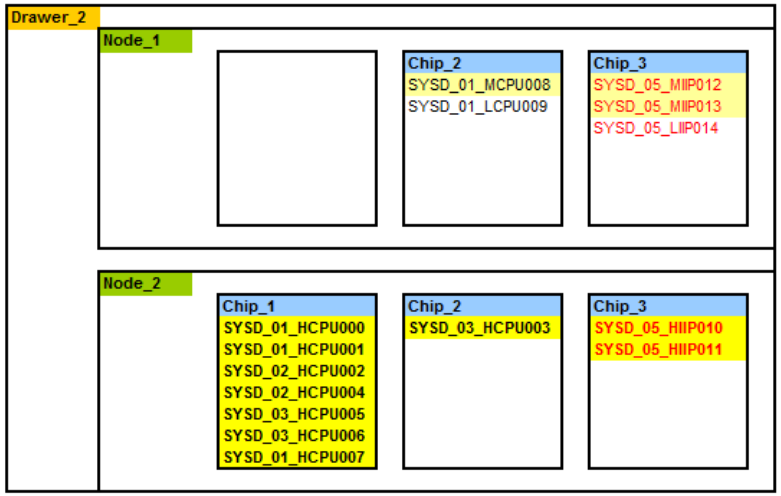
SMF 99 Subtype 14 – HiperDispatch Topology

- SMF 99 Subtype 14 contains HiperDispatch Topology data including:
  - Logical Processor characteristics: Polarization (VH, VM, VL), Affinity Node, etc.
  - Physical topology information
    - zEC12 Book / Chip
    - z13 Drawer / Node / Chip
- Written every 5 minutes or when a Topology change occurs
    - e.g. Configuration change or weight change
- May be useful to help understand why performance changed
- Provides a “Topology Change” indicator
  - Can identify when the topology changed occurred
- Recommendation is to collect SMF 99 subtype 14s for each System / LPAR
- New *WLM Topology Report* available to process SMF 99 subtype 14 records
  - http://www.ibm.com/systems/z/os/zos/features/wlm/WLM_Further_Info_Tools.html#Topology

z13 SYSD *WLM Topology Report* – Jan 30

Topology before SYSD Tests

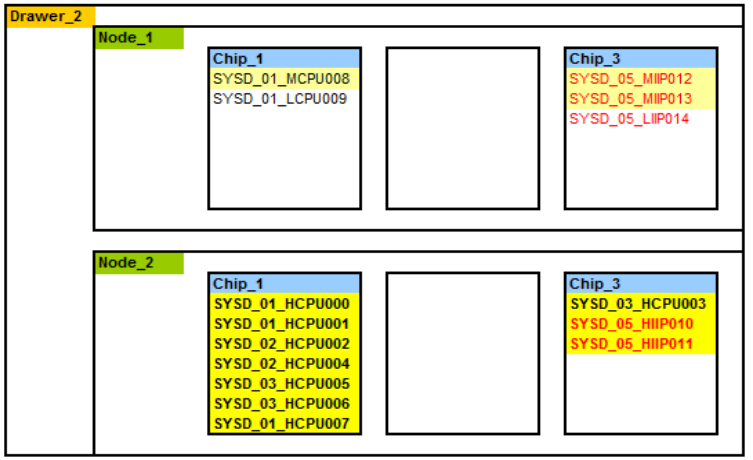
Topology for 01/30/2015-14:08:32, System: SYSD



Topology for 14:20 - 14:40 SYSD Tests

Changed at 14:11:42. Due to adding zIIPs on SYSB

Topology for 01/30/2015-14:11:42, System: SYSD



z13 SMT Capacity and Performance Metrics

WSC z13 zIIP SMT Test

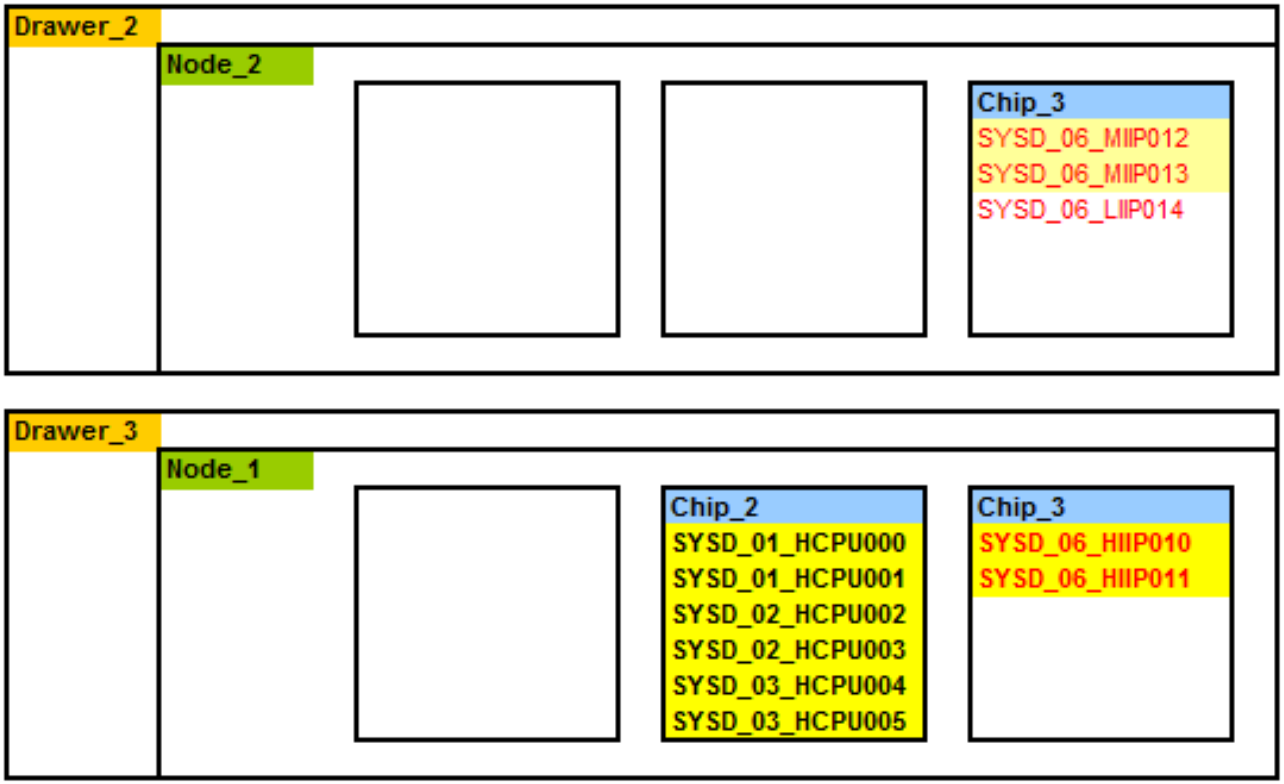
- z13- 736 N96 with 12 zIIPs
  - Processor at pre GA code level
  - Tests ran on partition SYSD / USP02 defined with 6 GCPs and 5 zIIPs
    - 5 zIIPS: 2 VH, 2 VM and 1 VL
  - Other partitions running very limited load, SYSD weights not enforced
  - z/OS 2.1
- External Server driving 2 Java 8 workloads
  - Gets memory and calculates PI to 6 digits (CB_MED)
  - Creates 25 page PDF document (CB_LOW)
  - Artificial driver to drive large number of transactions in 5 minute interval
  - Test consisted of 2 RMF/SMF 5 minute intervals
  - SMT-2 Vs SMT-1 with IIPHONORPRIORITY=YES and NO
- Objective
  - Utilize ~58% 5 Logical zIIPs and ~25% 12 Physical zIIPs
  - Observe Response Times and zIIP utilization with SMT-1 and SMT-2

These numbers come from a synthetic Benchmark and do not represent a production workload

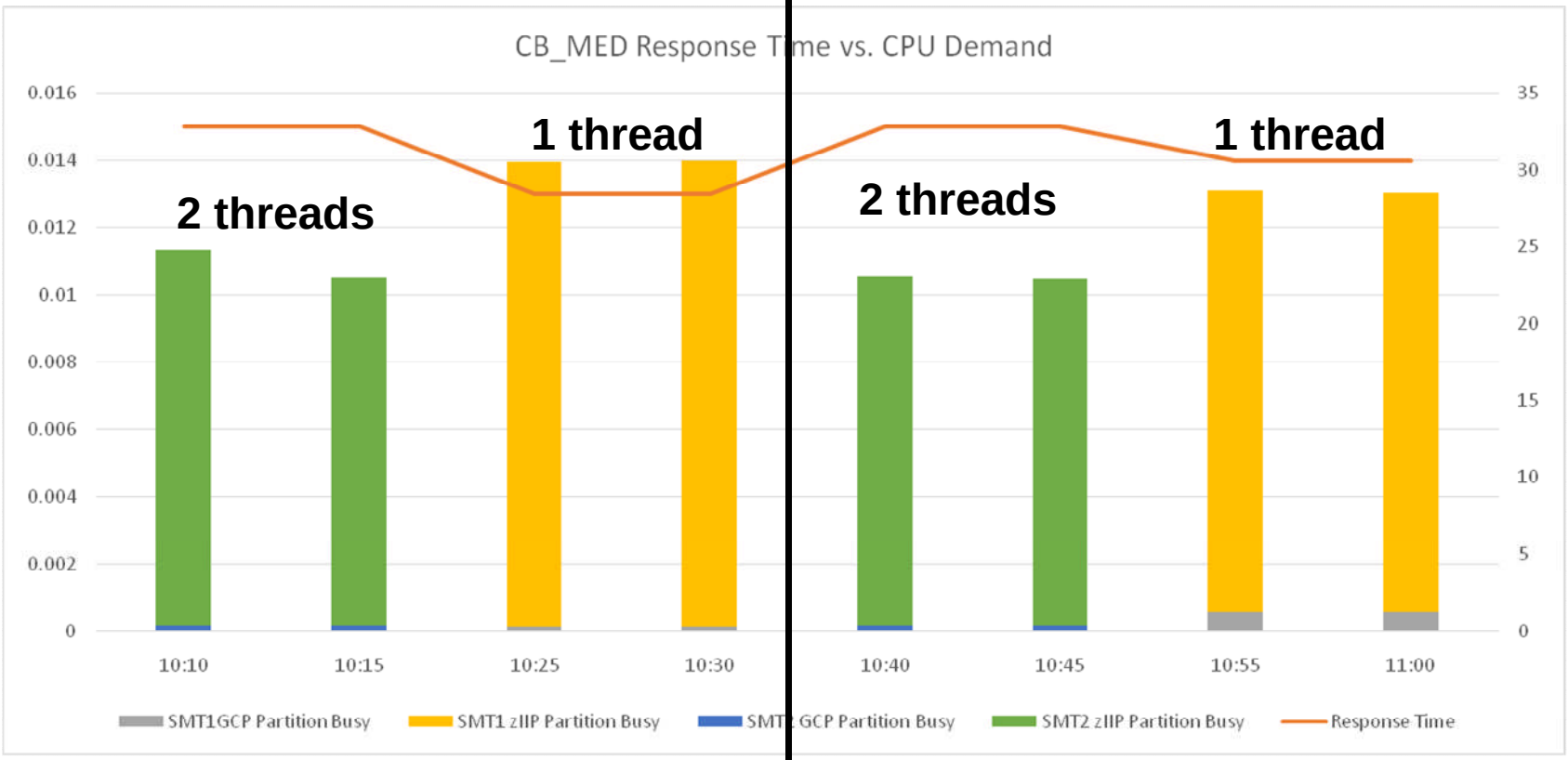
z13 SYSD *WLM Topology Report* – Feb 25 Tests

Topology for 02/25/2015-10:13:16 ,System: SYSD

2 zIIP VMs and 1 zIIP VL on Drawer 2, Node 2, Chip 3



WSC z13 zIIP SMT Test Summary



HonorPriority=NO

HonorPriority=YES

SMT-2 Mode resulted in

- **Lower zIIP Utilization**
- **CB_MED Higher Response times**

These numbers come from a synthetic Benchmark and do not represent a production workload

WSC z13 zIIP SMT Test Summary – CPU MF

CPU MF – zIIP Pool

RMF Start	SMT	TDEN	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L2P	L3P	L4LP	L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Machine Type	LSPR Wkld
10:10	2	1.79	1.92	95.9	1.47	0.45	40	1.1	74.6	13.8	4.7	2.8	4.1	1.40	478.6	5.0	Z13	AVG
10:15	2	1.75	1.93	95.4	1.49	0.44	40	1.1	73.8	13.6	5.2	2.9	4.5	1.50	429.8	5.0	Z13	AVG
10:25	1	1.00	1.63	95.4	1.15	0.48	52	0.9	67.6	16.5	6.7	3.7	5.6	1.87	359.7	5.0	Z13	AVG
10:30	1	1.00	1.64	95.4	1.15	0.49	52	0.9	67.4	16.7	6.7	3.7	5.5	1.86	361.2	5.0	Z13	AVG
10:40	2	1.74	1.93	95.4	1.49	0.44	41	1.1	73.9	13.5	5.2	2.9	4.6	1.51	427.4	5.0	Z13	AVG
10:45	2	1.74	1.93	95.3	1.49	0.44	41	1.1	73.7	13.6	5.3	2.9	4.5	1.51	427.4	5.0	Z13	AVG
10:55	1	1.00	1.62	95.3	1.17	0.45	48	0.9	69.6	14.8	6.9	3.3	5.4	1.79	326.8	5.0	Z13	AVG
11:00	1	1.00	1.61	95.4	1.17	0.45	48	0.9	69.8	14.7	6.9	3.3	5.4	1.78	325.4	5.0	Z13	AVG

SMT-1 has a lower CPI (faster) than SMT-2

These numbers come from a synthetic Benchmark and do not represent a production workload

SYSD RMF CPU Activity – zIIPs Feb 25 10:45 AM

C P U A C T I V I T Y

z/OS V2R1 SYSTEM ID SYSD DATE 02/25/2015
RPT VERSION V2R1 RMF TIME 10.45.00

CPU 2964 CPC CAPACITY 4452 SEQUENCE CODE 0000000000008DA97
MODEL 736 CHANGE REASON=NONE HIPERDISPATCH=YES
H/W MODEL N96

---CPU---		TIME %				--- MT % ---		LOG PROC	
NUM	TYPE	ONLINE	LPAR BUSY	MVS BUSY	PARKED	PROD	UTIL	SHARE %	
A	IIP	100.00	88.46	75.43	0.00	89.45	79.13	100.0	HIGH
				70.30	0.00				
B	IIP	100.00	81.96	68.84	0.00	87.79	71.95	100.0	HIGH
				59.34	0.00				
C	IIP	100.00	59.18	49.67	0.00	88.09	52.13	53.2	MED
				41.59	0.00				
D	IIP	100.00	31.78	27.10	0.00	88.09	27.99	53.2	MED
				22.31	0.00				
E	IIP	100.00	11.13	9.68	0.00	91.11	10.14	0.0	LOW
				8.33	0.00				
TOTAL/AVERAGE			54.50	43.26		88.91	48.27	306.4	

MULTI-THREADING ANALYSIS

CPU	TYPE	MODE	MAX CF	CF	AVG TD
	CP	1	1.000	1.000	1.000
	IIP	2	1.384	1.225	1.585

Max Capacity Factor (MAX CF) – How much work a core can complete (rate of delivery)
Capacity Factor (CF) – How much work a core actually completes (rate of delivery)
Average Thread Density (AVG TD) – Average executing threads during Core Busy
MT % Productivity (PROD) – Core Busy Time Effectiveness (Capacity in use / Capacity max)
MT % Utilization (UTIL) – Core Busy Time / Core Available Time

SYSD RMF CPU Activity – CPU MF Perspective

```
                                                C P U   A C T I V I T Y

            z/OS V2R1                  SYSTEM ID SYSD             DATE 02/25/2015
                                       RPT VERSION V2R1 RMF       TIME 10.45.00
CPU         2964   CPC CAPACITY  4452          SEQUENCE CODE 000000000008DA97
MODEL       736    CHANGE REASON=NONE          HIPERDISPATCH=YES
H/W MODEL   N96
---CPU---    ---------------- TIME % ----------------    --- MT % ----    LOG PROC
NUM  TYPE     ONLINE     LPAR BUSY    MVS BUSY    PARKED   PROD   UTIL     SHARE %

 A     IIP    100.00     88.46        75.43         0.00    89.45  79.13    100.0  HIGH
                                      70.30         0.00
 B     IIP    100.00     81.96        68.84         0.00    87.79  71.95    100.0  HIGH
                                      59.34         0.00
 C     IIP    100.00     59.18        49.67         0.00    88.09  52.13     53.2  MED
                                      41.59         0.00
 D     IIP    100.00     31.78        27.10         0.00    88.09  27.99     53.2  MED
                                      22.31         0.00
 E     IIP    100.00     11.13         9.68         0.00    91.11  10.14      0.0  LOW
                                       8.33         0.00
TOTAL/AVERAGE             54.50        43.26                 88.91  48.27    306.4
------------ MULTI-THREADING ANALYSIS ---------------
 CPU TYPE     MODE      MAX CF           CF       AVG TD
       CP        1       1.000        1.000        1.000
       IIP       2       1.384        1.225        1.585
```

“CPU MF” perspective:

- **MAX CF = 1 / CPI “Max” capacity for the mix**
- **CF = 1 / CPI “Actual” capacity for the mix**
- **TD = “Queue Length” >=1 and <=2 threads when busy**
- **PROD = CF / MAX CF**
- **UTIL = LPARBUSY “How Much”**

“MAX CF” estimate from a CPU MF perspective

- **If CPI MT-1 = 1.61, then IPC are 1 / 1.61 = .6211**
- **If CPI MT-2 = 2.33, then IPC are 1 / 2.33 = .4298**
- **MAX CF = (2 threads x .4298) / .6211 = 1.384**

capacity is rate of delivery in Instructions Per Cycle (IPC) for the workload mix

SYSD CPU MF Thread Metrics – zIIPs Feb 25 10:45 AM

```
                                C P U   A C T I V I T Y

            z/OS V2R1                  SYSTEM ID SYSD            DATE 02/25/2015
                                       RPT VERSION V2R1 RMF      TIME 10.45.00
CPU         2964   CPC CAPACITY  4452        SEQUENCE CODE 000000000008DA97
MODEL       736    CHANGE REASON=NONE        HIPERDISPATCH=YES
H/W MODEL   N96
---CPU---    ---------------- TIME % ----------------   --- MT % ----   LOG PROC
NUM  TYPE     ONLINE     LPAR BUSY    MVS BUSY   PARKED   PROD   UTIL     SHARE %

 A     IIP    100.00     88.46         75.43        0.00    89.45  79.13   100.0  HIGH
                                        70.30        0.00
 B     IIP    100.00     81.96         68.84        0.00    87.79  71.95   100.0  HIGH
                                        59.34        0.00
 C     IIP    100.00     59.18         49.67        0.00    88.09  52.13    53.2  MED
                                        41.59        0.00
 D     IIP    100.00     31.78         27.10        0.00    88.09  27.99    53.2  MED
                                        22.31        0.00
 E     IIP    100.00     11.13          9.68        0.00    91.11  10.14     0.0  LOW
                                         8.33        0.00
TOTAL/AVERAGE             54.50         43.26                88.91  48.27   306.4
------------ MULTI-THREADING ANALYSIS ---------------
 CPU TYPE     MODE      MAX CF           CF        AVG TD
      CP         1       1.000        1.000         1.000
      IIP        2       1.384        1.225         1.585
```

CPU MF - zIIPs

RMF Start	SMT	CPID	THREAD	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L2P	L3P	L4LP	L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Machine Type	LSPR Wkld	Drawer	Node	Chip	Logical
10:45	2	20	1	1.90	95.3	1.49	0.41	38	1.1	74.0	13.3	4.9	3.1	4.7	1.54	74.6	5.0	Z13	AVG	3	1	3	VH
10:45	2	21	2	2.03	94.9	1.60	0.43	38	1.1	73.9	13.3	4.9	3.1	4.8	1.57	69.5	5.0	Z13	AVG	3	1	3	VH
10:45	2	22	1	1.81	95.6	1.42	0.39	37	1.0	74.0	13.5	4.9	3.1	4.6	1.51	68.1	5.0	Z13	AVG	3	1	3	VH
10:45	2	23	2	1.99	94.9	1.56	0.43	38	1.1	73.8	13.6	4.9	3.1	4.6	1.53	58.7	5.0	Z13	AVG	3	1	3	VH
10:45	2	24	1	1.84	95.7	1.38	0.46	45	1.0	73.9	13.3	5.9	2.6	4.3	1.46	49.0	5.0	Z13	AVG	2	2	3	VM
10:45	2	25	2	2.08	95.0	1.58	0.50	45	1.1	73.6	13.6	5.9	2.5	4.4	1.48	41.0	5.0	Z13	AVG	2	2	3	VM
10:45	2	26	1	1.84	95.8	1.37	0.47	46	1.0	72.9	14.5	5.9	2.6	4.1	1.44	26.7	5.0	Z13	AVG	2	2	3	VM
10:45	2	27	2	2.05	95.2	1.54	0.51	46	1.1	72.8	14.6	5.8	2.6	4.3	1.46	22.0	5.0	Z13	AVG	2	2	3	VM
10:45	2	28	1	1.83	96.1	1.33	0.50	49	1.0	71.5	16.0	5.6	3.0	4.0	1.45	9.6	5.0	Z13	AVG	2	2	3	VL
10:45	2	29	2	2.07	95.5	1.51	0.56	51	1.1	71.1	15.9	5.8	3.0	4.2	1.50	8.2	5.0	Z13	AVG	2	2	3	VL
		Total		1.93	95.3	1.49	0.44	41	1.1	73.7	13.6	5.3	2.9	4.5	1.51	427.4	5.0	Z13	AVG				

SYSD CPU MF Thread Metrics – zIIPs Feb 25 SMT-2 Vs SMT-1

CPU MF - zIIPs

PROCVIEW CORE, MT_ZIIP_MODE=2

RMF Start	SMT	CPID	THREAD	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L2P	L3P	L4LP	L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Machine Type	LSPR Wkld	Drawer	Node	Chip	Logical
10:45	2	20	1	1.90	95.3	1.49	0.41	38	1.1	74.0	13.3	4.9	3.1	4.7	1.54	74.6	5.0	Z13	AVG	3	1	3	VH
10:45	2	21	2	2.03	94.9	1.60	0.43	38	1.1	73.9	13.3	4.9	3.1	4.8	1.57	69.5	5.0	Z13	AVG	3	1	3	VH
10:45	2	22	1	1.81	95.6	1.42	0.39	37	1.0	74.0	13.5	4.9	3.1	4.6	1.51	68.1	5.0	Z13	AVG	3	1	3	VH
10:45	2	23	2	1.99	94.9	1.56	0.43	38	1.1	73.8	13.6	4.9	3.1	4.6	1.53	58.7	5.0	Z13	AVG	3	1	3	VH
10:45	2	24	1	1.84	95.7	1.38	0.46	45	1.0	73.9	13.3	5.9	2.6	4.3	1.46	49.0	5.0	Z13	AVG	2	2	3	VM
10:45	2	25	2	2.08	95.0	1.58	0.50	45	1.1	73.6	13.6	5.9	2.5	4.4	1.48	41.0	5.0	Z13	AVG	2	2	3	VM
10:45	2	26	1	1.84	95.8	1.37	0.47	46	1.0	72.9	14.5	5.9	2.6	4.1	1.44	26.7	5.0	Z13	AVG	2	2	3	VM
10:45	2	27	2	2.05	95.2	1.54	0.51	46	1.1	72.8	14.6	5.8	2.6	4.3	1.46	22.0	5.0	Z13	AVG	2	2	3	VM
10:45	2	28	1	1.83	96.1	1.33	0.50	49	1.0	71.5	16.0	5.6	3.0	4.0	1.45	9.6	5.0	Z13	AVG	2	2	3	VL
10:45	2	29	2	2.07	95.5	1.51	0.56	51	1.1	71.1	15.9	5.8	3.0	4.2	1.50	8.2	5.0	Z13	AVG	2	2	3	VL
			Total	1.93	95.3	1.49	0.44	41	1.1	73.7	13.6	5.3	2.9	4.5	1.51	427.4	5.0	Z13	AVG				

PROCVIEW CORE, MT_ZIIP_MODE=1

RMF Start	SMT	CPID	THREAD	CPI	Prb State	Est Instr Cmplx CPI	Est Finite CPI	Est SCPL1M	L1MP	L2P	L3P	L4LP	L4RP	MEMP	Rel Nest Intensity	LPARCPU	Eff GHz	Machine Type	LSPR Wkld	Drawer	Node	Chip	Logical
11:00	1	20	1	1.57	95.1	1.16	0.41	44	0.9	70.2	14.2	7.4	3.0	5.3	1.76	75.7	5.0	Z13	AVG	3	1	3	VH
11:00	1	22	1	1.58	95.1	1.17	0.41	44	0.9	70.0	14.3	7.4	3.0	5.3	1.76	73.0	5.0	Z13	AVG	3	1	3	VH
11:00	1	24	1	1.66	95.4	1.18	0.48	52	0.9	69.9	14.6	6.6	3.5	5.5	1.80	68.3	5.0	Z13	AVG	2	2	3	VM
11:00	1	26	1	1.64	95.6	1.17	0.47	52	0.9	69.6	15.1	6.4	3.5	5.4	1.79	59.9	5.0	Z13	AVG	2	2	3	VM
11:00	1	28	1	1.64	95.9	1.16	0.48	53	0.9	68.8	15.9	6.3	3.6	5.4	1.81	48.5	5.0	Z13	AVG	2	2	3	VL
			Total	1.61	95.4	1.17	0.45	48	0.9	69.8	14.7	6.9	3.3	5.4	1.78	325.4	5.0	Z13	AVG				

Looking for z13 Migration "Volunteers" to send SMF data

- Want to validate / refine Workload selection metrics

Looking for "Volunteers"

(3 days, 24 hours/day, SMF 30s, 70s, 72s, 99 subtype 14s,113s per LPAR)

"Before z196 / zEC12" and "After z13"

Production partitions preferred

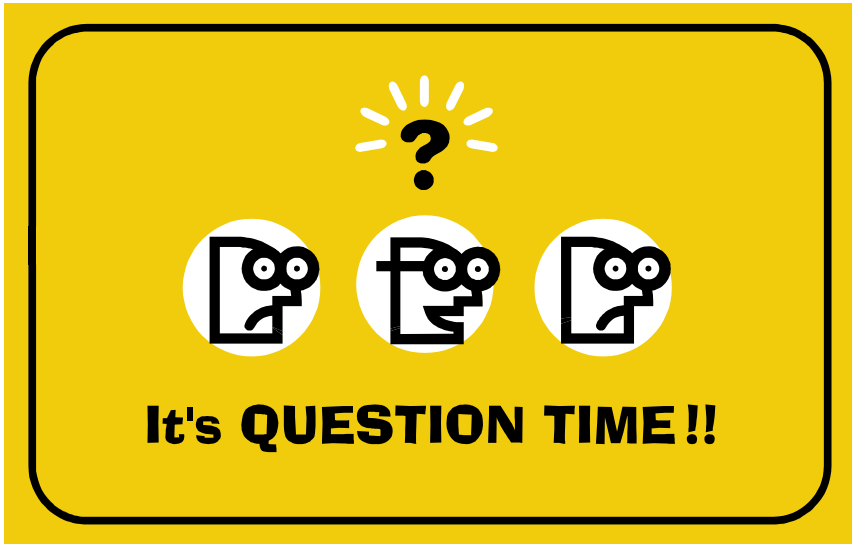
If interested send note to jpburg@us.ibm.com,

No deliverable will be returned

Benefit: Opportunity to ensure your data is used to influence analysis

CPU MF 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 all your systems



Thank You for Attending!

Techdocs provides the latest ATS technical collateral

www.ibm.com/support/techdocs

Back Up

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**

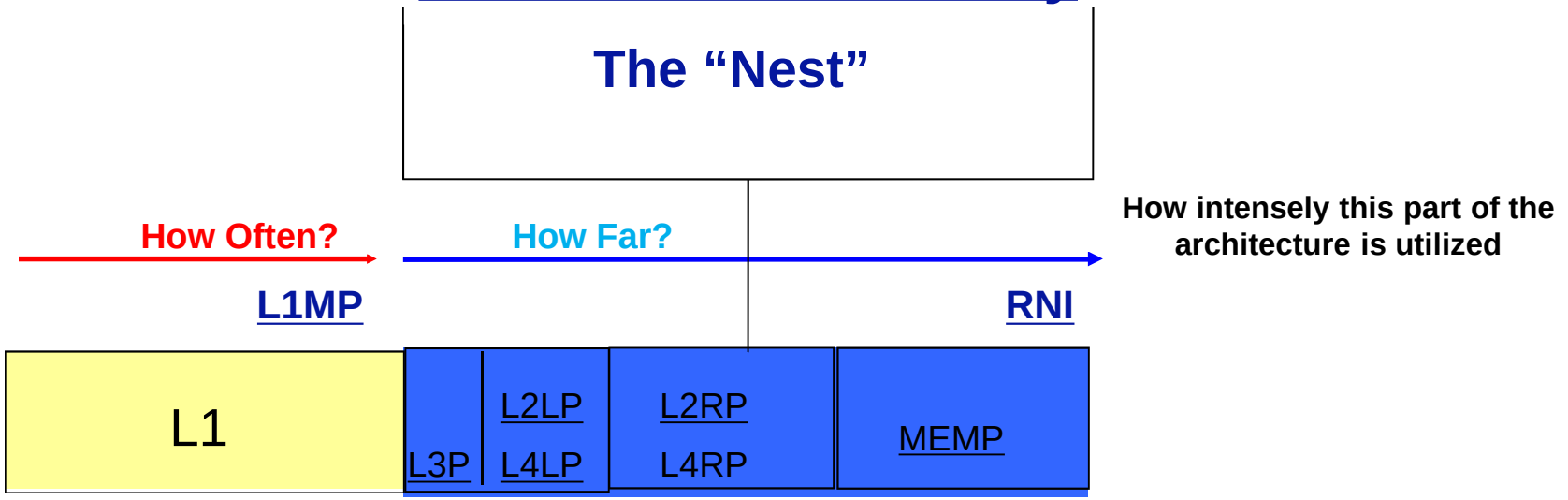
Workload Capacity Performance

- 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**
  - **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**
  - **For zEC12 / zBC12 RNI = 2.3 *(0.4*L3P + 1.2*L4LP + 2.7*L4RP + 8.2*MEMP) / 100**
  - **For z13 RNI = 2.6 *(0.4*L3P + 1.6*L4LP + 3.5*L4RP + 7.5*MEMP) / 100**

Relative Nest Intensity



Microprocessor Design **Memory Hierarchy or Nest**

* zEC12 / zBC12 RNI Changed January 2015

*z196 / z114 RNI Changed July 2012

Note these Formulas may change in the future

Definitions

CPI – Cycles per Instruction

PRB STATE - % Problem State

L1MP – Level 1 Miss Per 100 instructions

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

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

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

L3P – % sourced from L3 cache

MEMP - % sourced from Memory

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

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

Rel Nest Intensity – Relative Nest Intensity

Eff GHz – Effective Gigahertz

Machine Type – Machine Type (e.g. z10, z196, zEC12)

LSPR Wkld – LSPR Workload match based on L1MP and RNI

Pool – 1 = GCP, 3 = zAAP, 6 = zIIP