Refresher of SMF 113 Processor Cache Counters

Peter Enrico

Email: Peter.Enrico@EPStrategies.com

Enterprise Performance Strategies, Inc.
3457-53rd Avenue North, #145
Bradenton, FL 34210
http://www.epstrategies.com
http://www.pivotor.com

Voice: 813-435-2297
Mobile: 941-685-6789

z/OS Performance Education, Software, and Managed Service Providers

Creators of Pivotor®

Contact, Copyright, and Trademark Notices

Questions?

Send email to Peter at Peter.Enrico@EPStrategies.com, or visit our website at http://www.epstrategies.com or http://www.pivotor.com.

Copyright Notice:

Trademarks:

Enterprise Performance Strategies, Inc. presentation materials contain trademarks and registered trademarks of several companies.

The following are trademarks of Enterprise Performance Strategies, Inc.: **Health Check®, Reductions®, Pivotor®**

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries: IBM®, z/OS®, zSeries® WebSphere®, CICS®, DB2®, S390®, WebSphere Application Server®, and many others.

Other trademarks and registered trademarks may exist in this presentation

Abstract

- **Refresher of SMF 113 Processor Cache Counters**

  - The SMF 113 record has been around for while now. It is expected that a new processor will be announced by IBM sometime in, the somewhat, near future. Understanding, evaluating, and using the SMF 113 processor cache measurements is always critical when evaluating your current processor or considering an upgrade.

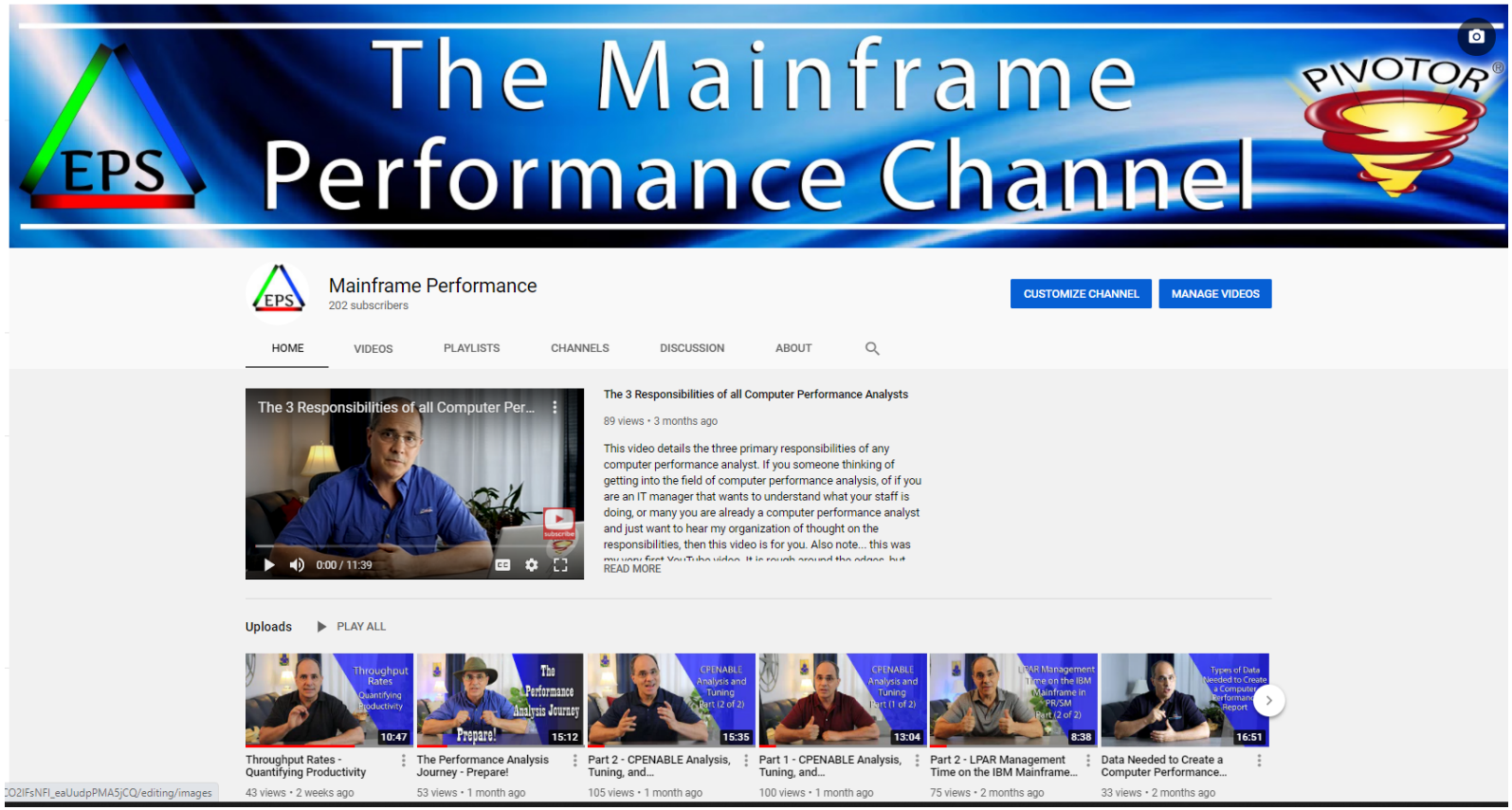
  - During this presentation, Peter Enrico will provide a back-to-basics presentation with the objective of reminding everyone why the SMF 113 processor cache measurements are interesting for analysis.

EPS: We do z/OS performance…

- We love to bring you these complimentary webinars and to teach you how to do what we do!

- But remember, we would love to work with you!
  - ***Pivotor - Performance reporting and analysis of your z/OS measurements (i.e. SMF, etc.)***
    - Not just reporting, but coset effective analysis-based reporting based on our expertise
  - ***Performance Educational Workshops (while analyzing your own data)***
    - Essential z/OS Performance Tuning
    - Parallel Sysplex and z/OS Performance Tuning
    - WLM Performance and Re-evaluating Goals
  - ***Performance War Rooms***
    - Concentrated, highly productive group discussions and analysis

Subscribe YouTube Mainframe Performance Channel



Like what you see?

- Free z/OS Performance Educational webinars!
  - The titles for our Fall 2020 webinars are as follows:
    - ✓ *Advantages of Multiple Period Service Classes*
    - ✓ *Refresher of SMF 113 Processor Cache Counters and Concepts*
    - *WLM SYSTEM / SYSSTC*
    - *Using Long Term Reporting: Pivotor Past Perfect*
    - *Catching Up with Bob Rogers*
      - *Specific Topic to be decided? Suggestions?*
    - *System Recovery Boost (SRB): The Turbo Button for z/OS*
    - *Data in Memory (DIM) Primer*
    - *ETR vs ITR*
    - *Inventory Those Managed Resources and Workloads*
    - *Data Visualization - Pivotor People Pontificate*
  - Let us know if you want to be on our mailing list for these webinars

- If you want a free cursory review of your environment, let us know!
  - We're always happy to process a day's worth of data and show you the results
  - See also: http://pivotor.com/cursoryReview.html

Why do we care about processor cache measurements and usage of the caching hierarchy?

Key Influences of Processor Performance and Capacity

- Question:

  What are the key influences that result in variations of a particular processor's delivered capacity relative to a customer's environment and workload?

- That is:
  - Why is it when you size a processor using zPCR or IBM's LSPR charts, your results may vary from zPCR or IBMs?
  - Why could the same machine's capacity vary based on a particular workload?
  - Why could the same machine's capacity vary based on a particular customer?
  - Why could the same machine's capacity vary based on a particular configuration setting in z/OS, WLM, PR/SM, etc?

Key Influences of Processor Performance and Capacity

- Question:

  What are the key influences that result in variations of a particular processor's delivered capacity relative to a customer's environment and workload?

- Answer: As Gary King of IBM would say… there are three key influences:
  - Instruction complexity of one processor family to another
  - Path length of the code executed by customer applications and transactions
  - Usage of the Memory Hierarchy
- A machine's capacity will vary based on each of these three factors

Key Influence - Instruction Complexity

- Instruction complexity of one processor family to another
  - Types of instructions, the sequence in which they are executed, and the way they interact with the processor design
  - Many machine design alternatives affect instruction complexity
  - Each processor family has variations in the chip design
  - As processor evolve the way the chip executes instructions is enhance and geared towards the technology
- Examples include cycle time of CPU, how the instructions are wired to execute (using pipelining, branch prediction, out of order execution, etc.)
- Influences the workload
  - Once a customer workloads are on a processor, instruction complexity is relatively constant between customers and workloads
  - In other words, relative to the LSPRs and sizing, once a move is made to the new processor family instruction complexity does not vary much from one customer to the next.

Key Influence - Instruction Complexity

- Notice the PCI column showing the PCIs of the 701 series of each processor
  - PCI : Processor Capacity Index

    Term used by IBM instead of MIPS

							Core-level			
zGen	Name	Year	Mach Type	GHz	701 PCI	701 MSUs	L1-Data	L1-Instr	L2-Data	L2-Instr
z9	z9 EC	2005	2094	1.7	560	81	256K	256K	n/a	n/a
z10	z10 EC	2008	2097	4.4	902	115	128K	64K	3M	
z11	z196	2010	2817	5.2	1202	150	128K	64K	1.5M	
z12	zEC12	2012	2827	5.5	1514	188	96K	64K	1M	1M
z13	z13	2015	2964	5	1695	210	128K	96K	2M	2M
z14	z14	2017	3906	5.2	1832	227	128K	128K	4M	2M
z15	z15	2019	8561	5.2	2055	253	128K	128K	4M	4M

zEC12 to z14:
Cycle time: -5.5%
Performance: +21%

z13 to z14:
Cycle time: +4%
Performance: +8%

z14 to z15:
Cycle time: +0%
Performance: +12%

Key Influence – Path Length

- Path length of the code executed by customer applications and transactions
  - This relates to code executed by applications / jobs / transactions / etc.
  - Instruction count
- The actual path lengths executed by a workload will vary
  - From customer to customer, and from IBM synthetic workloads versus customer
  - From one customer's application environment versus another application environment of that same customer
    - Example: CICS / DB2 application versus a WAS / DB2 application
- Is sensitive to the configuration due to MP effects
  - Higher n-ways or difference in configuration may increase path lengths execute (which in turn influences the processor capacity relative to LSPRs)
    - Example: May have more locking in a higher MP environment, or queues may be longer, etc.
- But when move from one processor to another this generally does not change much for a specific customer
  - Whether the move is from one processor family to another
  - Or from one process in the same family to another

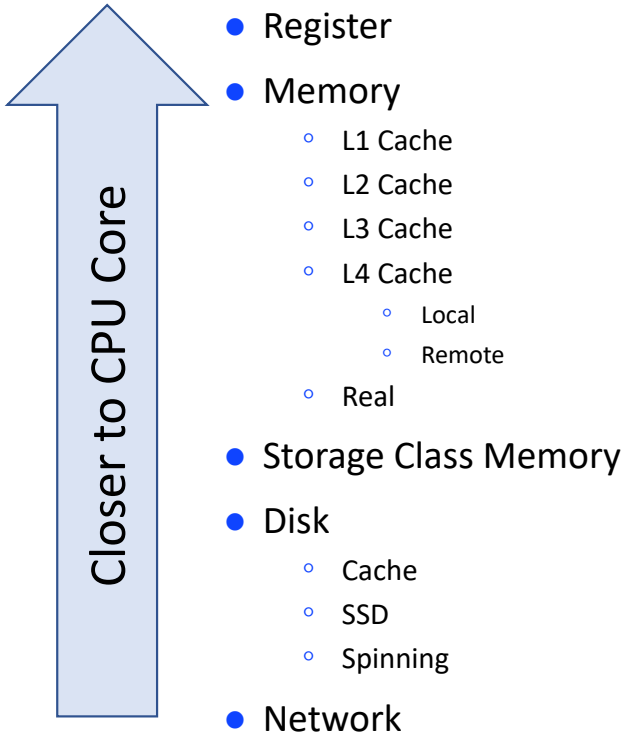
Key Influence – Memory Hierarchy

- Usage of the Memory Hierarchy
  - Heavily influenced by key factors result potentially wide variations in realized capacity
  - From one processor family to another there are many design alternatives
    - Levels of cache, scope of cache, latency, etc.
  - Configuration will influence usage of the memory hierarchy
    - LPAR configuration, competition between LPARs, options such as HiperDispatch, etc.
  - Exploitation by workloads will influence usage of the memory hierarchy
    - Transaction intensity, memory intensity, I/O intensity, application mixtures, competition of resource by applications, etc.
  - z/OS performance management and options
    - WLM management of resources, affinity nodes, IEAOPTxx opts, heap sizes, initiators, etc.
- Final result is that usage of memory hierarchy heavily influences a processor's delivered capacity and performance.
  - Workload performance sensitive to how deep into the memory hierarchy the processor must go to retrieve instructions and data
- So, for processor sizing, LSPRs have started focusing on this

Introducing the Processor Caches of IBM's zArchitecture Processors

Modern Performance Optimization

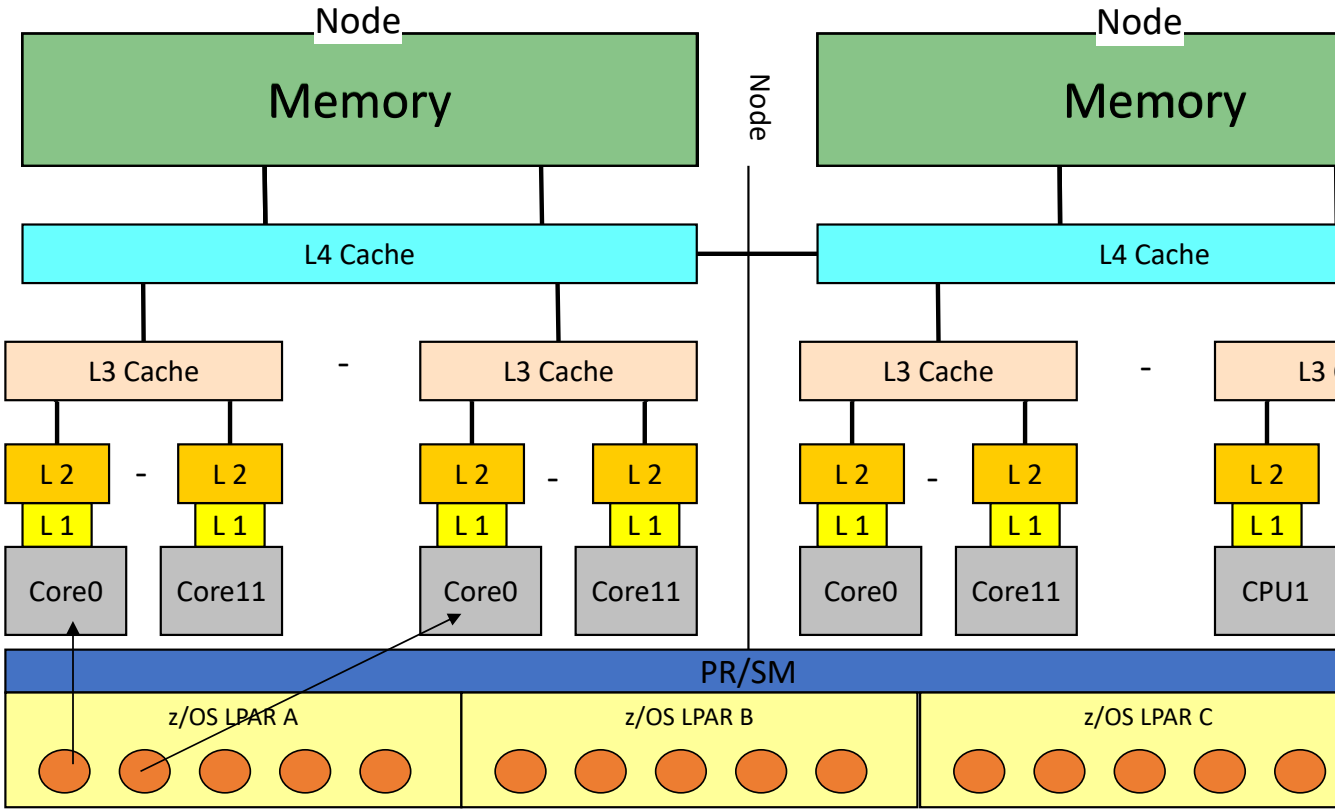
- Distance matters!
- Keep data close to not just the processor, but close to the instruction units on the processor
  - I.E. L1 cache hits very important
  - If the data isn't in L1, hopefully it's in L2, L3 or L4
- Hardware Instrumentation Services (HIS) records processor efficiency metrics in SMF 113 records
  - Be sure to record these
- SMF 99.14 records record mapping of logical to physical cores
  - Of particular interest for multi-book machines to make sure LPARs aren't crossing books



z Processor Cache Hierarchy

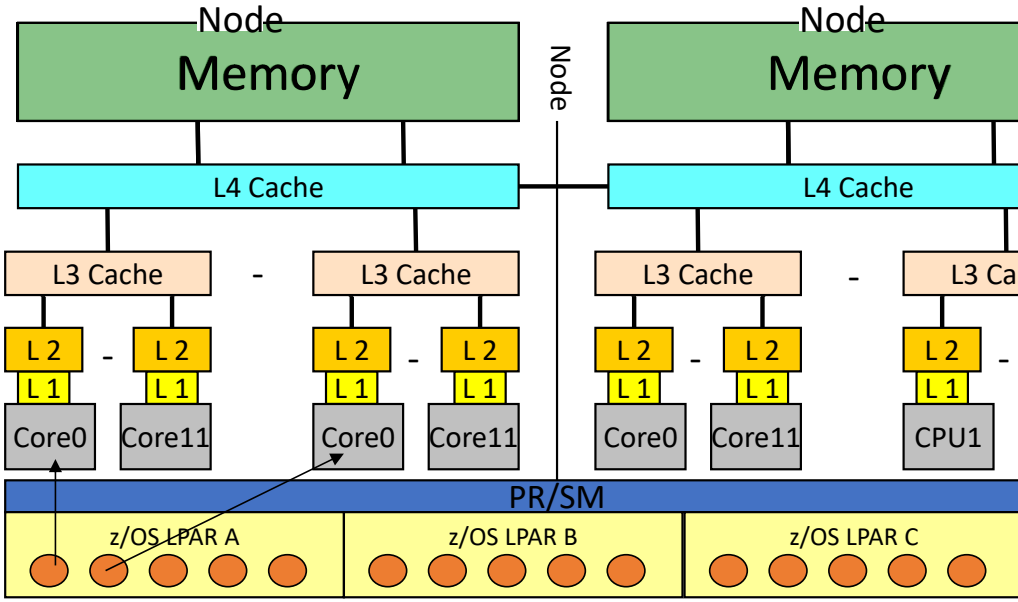
- Distance Matters!
  - Register
  - Processor Caches
    - L1 Cache
    - L2 Cache
    - L3 Cache
    - L4 Cache
      - Local
      - Remote
    - Real Memory (i.e. Central Storage)
  - Storage Class Memory
  - Disk
    - Cache
    - SSD
    - Spinning
  - Network

- Hardware Instrumentation Services (HIS) records processor efficiency metrics in SMF 113 records



Greatest Usage of SMF 113

- To understand the variability of a machine's capacity based on the usage of the processor cache memory hierarchy
  - IBM's LSPR Workloads
- Used to illustrate the usage of the processor caches to better understand before and after changes
  - Not good for benchmarking
  - But good to assuage concerns or gain insights
- Usage of standard SMF records still required for full processor evaluations
  - SMF 30
  - SMF 70
  - SMF 72.3
  - Etc..



LSPR Table Example – Pre SMF 113 (old way)

IBM System z10 EC (System z9 2094-701 = 1.00)

Processor	#CP	PCI	MSU	Mixed	LoIO-Mix	TI-Mix	CB-L	ODE-B	WASDB	OLTP-W	OLTP-T
2097-401	1	219	27	0.38	0.38	0.38	0.39	0.4	0.38	0.38	0.39
2097-402	2	414	51	0.73	0.73	0.72	0.76	0.77	0.73	0.69	0.73
2097-403	3	602	75	1.06	1.06	1.04	1.12	1.13	1.08	0.98	1.06
2097-404	4	782	97	1.37	1.39	1.34	1.47	1.48	1.41	1.25	1.37
2097-405	5	957	118	1.68	1.7	1.63	1.82	1.82	1.74	1.52	1.67
2097-406	6	1129	139	1.98	2.01	1.92	2.16	2.16	2.06	1.77	1.97
2097-407	7	1295	160	2.27	2.31	2.2	2.49	2.49	2.38	2.02	2.26
2097-408	8	1458	180	2.56	2.6	2.46	2.82	2.82	2.69	2.26	2.54
2097-409	9	1617	199	2.84	2.89	2.73	3.14	3.14	2.99	2.49	2.81
2097-410	10	1772	218	3.11	3.17	2.98	3.45	3.46	3.29	2.71	3.08
2097-411	11	1923	237	3.37	3.44	3.23	3.76	3.76	3.59	2.93	3.33
2097-412	12	2070	255	3.63	3.71	3.47	4.06	4.07	3.88	3.14	3.58
2097-501	1	473	58	0.83	0.83	0.83	0.85	0.86	0.82	0.81	0.83
2097-502	2	894	110	1.57	1.58	1.55	1.64	1.65	1.58	1.48	1.58
2097-503	3	1296	160	2.27	2.29	2.23	2.42	2.43	2.32	2.1	2.28
2097-504	4	1681	207	2.95	2.98	2.88	3.17	3.19	3.04	2.68	2.95
2097-505	5	2055	252	3.6	3.65	3.5	3.91	3.94	3.74	3.24	3.6
2097-506	6	2418	296	4.24	4.3	4.1	4.63	4.67	4.42	3.78	4.23

Reminder: 2097 = z10EC

LSPRs and SMF 113s and RNI Hint (a measure of usage of memory hierarchy)

- SMF 113 measurements are now used to provide guidelines / hints for LSPR and zPCR processor sizing
- This RNI Hint table was documented in the Large System Performance Reference (LSPR)
  - Google 'IBM LSPR'
- The next slide shows an example of an LSPR chart used for processor sizing
- Using the SMF 113 records you now need to calculate
  - L1MP - L1 Miss Per 100 Instructions
  - RNI – Relative Nest Intensity
- Note: This table and these guidelines are expected to change as more is learned from the SMF 113 records

L1MP	RNI	Workload Hint
<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

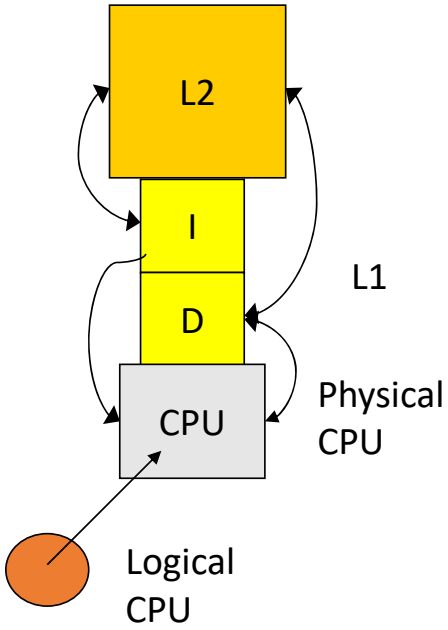
LSPR Table Example – Post SMF 113 (z14 LSPRs)

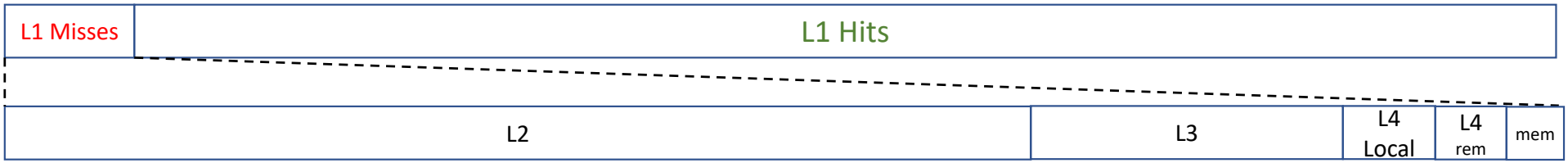
Processor	#CP	PCI**	MSU***	MSUps****	Low*	Average*	High*
3906-701	1	1832	227	182	3.37	3.27	3.04
3906-702	2	3464	427	342	6.55	6.19	5.61
3906-703	3	5050	620	496	9.66	9.02	8.08
3906-704	4	6590	808	646	12.68	11.77	10.45
3906-705	5	8082	990	792	15.64	14.44	12.74
3906-706	6	9528	1162	930	18.55	17.02	14.96
3906-707	7	10927	1326	1061	21.41	19.52	17.11
3906-708	8	12283	1487	1190	24.21	21.94	19.18
3906-709	9	13597	1642	1314	26.96	24.29	21.2
3906-710	10	14869	1793	1434	29.66	26.56	23.14
3906-711	11	16101	1939	1551	32.31	28.76	25.03
3906-712	12	17294	2077	1662	34.91	30.89	26.85
3906-713	13	18450	2213	1770	37.47	32.96	28.62
3906-714	14	19570	2346	1877	39.97	34.96	30.33
3906-715	15	20655	2476	1981	42.43	36.9	31.99

SMF 113 Counters

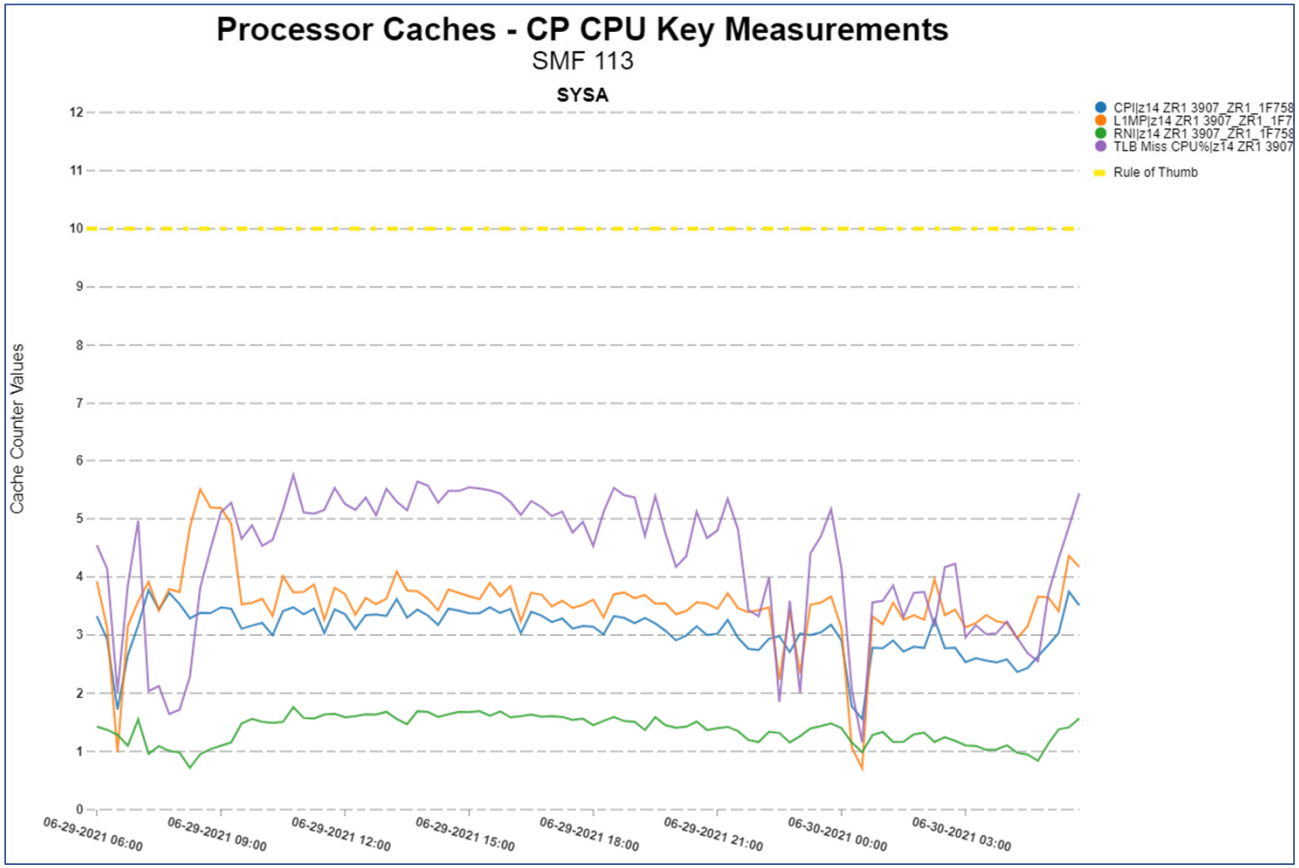
From these counters we have many useful calculated values

- Basic Counters
  - Used to understand the activity of the CPU and L1 cache
- Problem Counters
  - Problem state counters (subset of Basic Counters) of activity of the CPU and L1 cache
- Crypto Counters
  - Crypto processor function calls and blocks broken down by algorithm
- Extended Counters
  - Used to understand the 'sourcing' of L1 from L2, L3, L4 (local and remote), and memory





There are 4 key values (Pivotor Example)



4 key calculated (CP CPU example):

- CPI: Cycles per Instruction
  - Used to gauge processor contention, as well as instruction mix consistency
- L1MP: L1 misses per 100 instructions
  - Think of this as a 'miss percentage'
- RNI: Relative Next Intensity
  - Workload 'signature' that gauges the pressures being put on the upper caches and memory
- TLB Miss CPU Percentage
  - Total percent of the CPU consumed by the LPAR that goes to dynamic address translation (DAT) due to a translation look-aside buffer miss

Key SMF 113 Calculated Values

- CPI – Cycles Per Instruction
  - Simply calculated as number of cycles in interval / instructions completed
  - Estimated Finite CPI – CPI due to the fact that not all memory references are satisfied in L1 (i.e. because the L1 cache is finite)
    - Calculated via IBM formula (more directly on latest processors)
  - Instruction Complexity CPI – CPI due to the fact that some instructions simply take longer than others to execute
    - Calculated as CPI – Estimated Finite CPI

- L1MP – Level 1 Misses per 100 Instructions
  - Gives you an indication of how well you're leveraging L1 cache
  - Generally expect to be under 5 in most cases

Key SMF 113 Calculated Values

- Relative Nest Intensity
  - IBM formula, changes occasionally as new information becomes available about how the processors are actually performing in the field
  - See http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TC000066
- TLB CPU Miss Percent of CPU
  - Total percent of the CPU consumed by the LPAR that goes to dynamic address translation (DAT) due to a translation look-aside buffer miss
  - DAT is more costly than you might imagine: hope for it to be less than 5%, but not unusual for it to be more
  - Except for z14+: TLB redesign basically includes the DAT for every L1 cache line
    - 1-3% seems to be common for z14 and later
- For all these metrics, best to look at the metrics on a GCP vs. zIIP basis
  - Workload and utilization differences between the processor types result in differences in the metrics, averaging them together skews the metrics

Comments from Jamie…
and then
Q & A

Questions about content of webinar?

Of maybe general performance questions?

Like what you see?

- Free z/OS Performance Educational webinars!
  - The titles for our Fall 2020 webinars are as follows:
    - ✓ *Advantages of Multiple Period Service Classes*
    - ✓ *Refresher of SMF 113 Processor Cache Counters and Concepts*
    - *WLM SYSTEM / SYSSTC*
    - *Using Long Term Reporting: Pivotor Past Perfect*
    - *Catching Up with Bob Rogers*
      - *Specific Topic to be decided? Suggestions?*
    - *System Recovery Boost (SRB): The Turbo Button for z/OS*
    - *Data in Memory (DIM) Primer*
    - *ETR vs ITR*
    - *Inventory Those Managed Resources and Workloads*
    - *Data Visualization - Pivotor People Pontificate*
  - Let us know if you want to be on our mailing list for these webinars

- If you want a free cursory review of your environment, let us know!
  - We're always happy to process a day's worth of data and show you the results
  - See also: http://pivotor.com/cursoryReview.html