

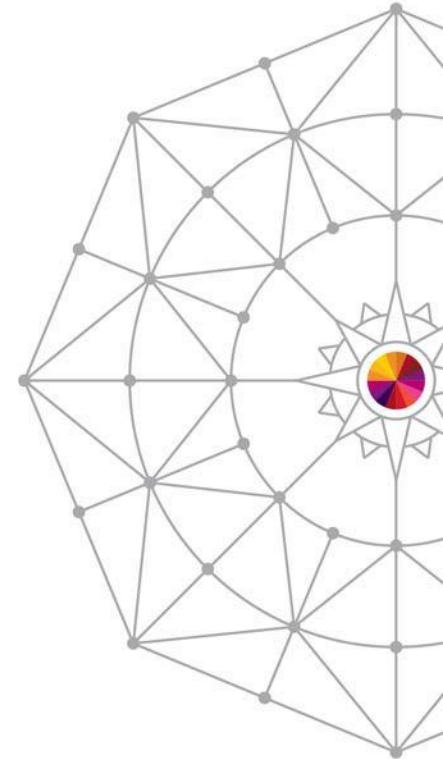
# Why is the CPU Time For a Job so Variable?

*Cheryl Watson, Frank Kyne  
Watson & Walker, Inc.*

[www.watsonwalker.com](http://www.watsonwalker.com)

[technical@watsonwalker.com](mailto:technical@watsonwalker.com)

*August 5, 2014, Session 15836*



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

Copyright (c) 2014 by SHARE Inc.

Except where otherwise noted, this work is licensed under <http://creativecommons.org/licenses/by-nc-sa/3.0/>

Except where otherwise noted, this work is licensed under <http://creativecommons.org/licenses/by-nc-sa/3.0/>

# Abstract



You run a job one day and it takes 3 CPU seconds and the next day it takes 5 seconds, and you didn't change anything. What happened? What can you do about it? Your billing and accounting is "screwed" up. The outsourcers and their customers are screaming at one another. What's going on?

Cheryl Watson and Frank Kyne, who have been watching this problem grow exponentially for years, have some answers as to what has happened and what you can do about it.

A good introduction to this session is the free SHARE webinar from Cheryl on "The Many CPU Fields of SMF" or Cheryl's previous SHARE presentations with the same title. Also available on our website under 'Presentations' at [www.watsonwalker.com](http://www.watsonwalker.com).



# Why is the CPU Time For a Job so Variable?

Watson & Walker:

- *Cheryl Watson's Tuning Letter*
- *Cheryl Watson's System z CPU Charts*
- Software products BoxScore & GoalTender
- Consulting
- Classes
- z/OS advocates

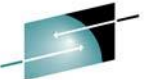


# Agenda

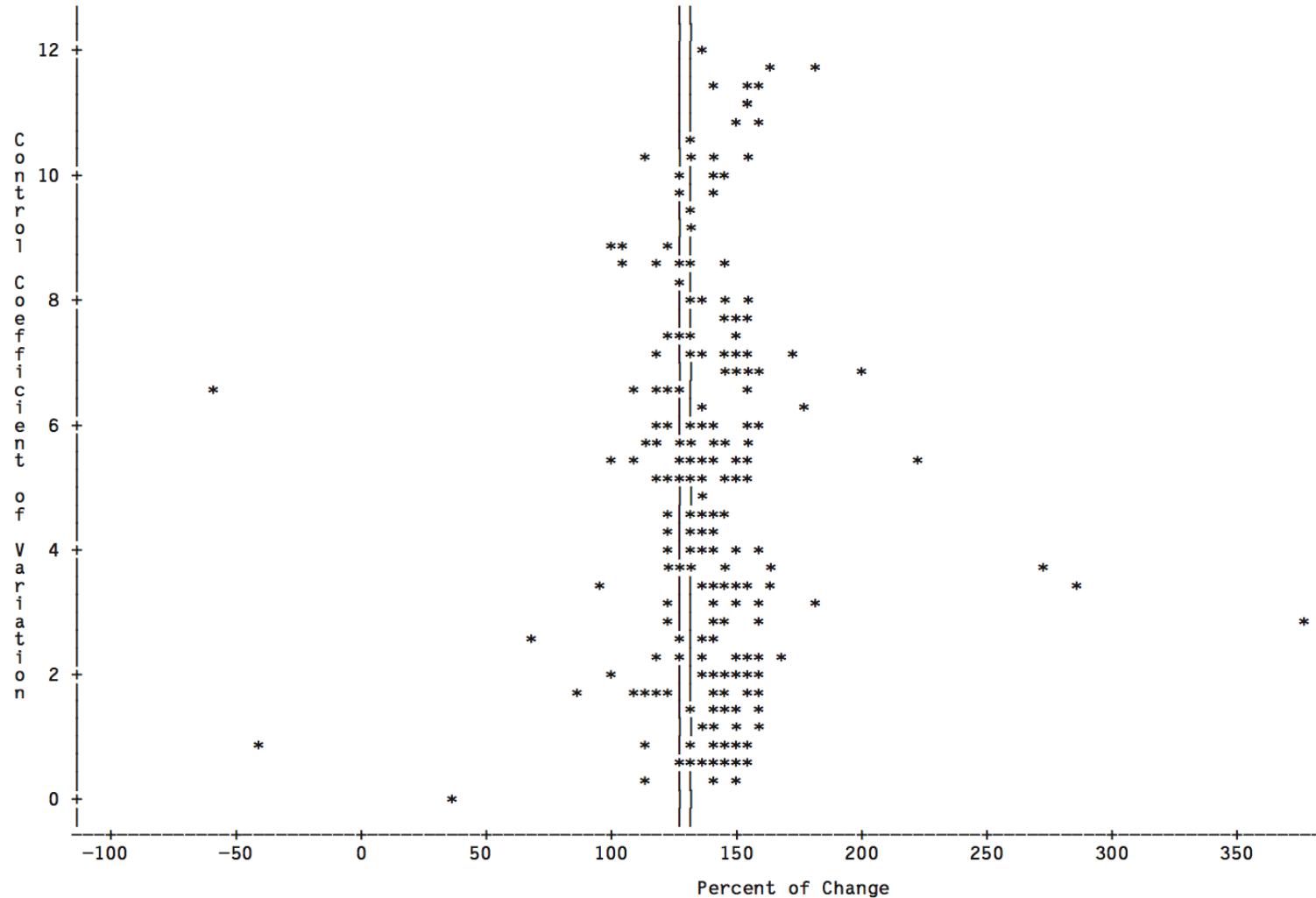
- CPU per I/O
  - Why This Topic?
    - Recent hardware changes
    - Recent customer experiences
  - Hardware Changes Affecting Variability
  - z/OS Changes Affecting Variability
  - Environment Changes Affecting Variability
  - Other Changes Affecting Variability
  - Which Measurements?
  - What To Do?

# CPU per I/O

- Ways to benchmark jobs:
  - Ideally, run at 90% or 99% busy, with all other work being the same (like IBM uses to come up with ITRRs)
  - Select jobs that run after every change using identical data; problem occurs because of other load on the system
  - Best, find every stable job step and look at the change in the CPU per I/O (take total CPU time and divide by the number of EXCPs) between two environments; this allows you to find new work that is affected by any type of change

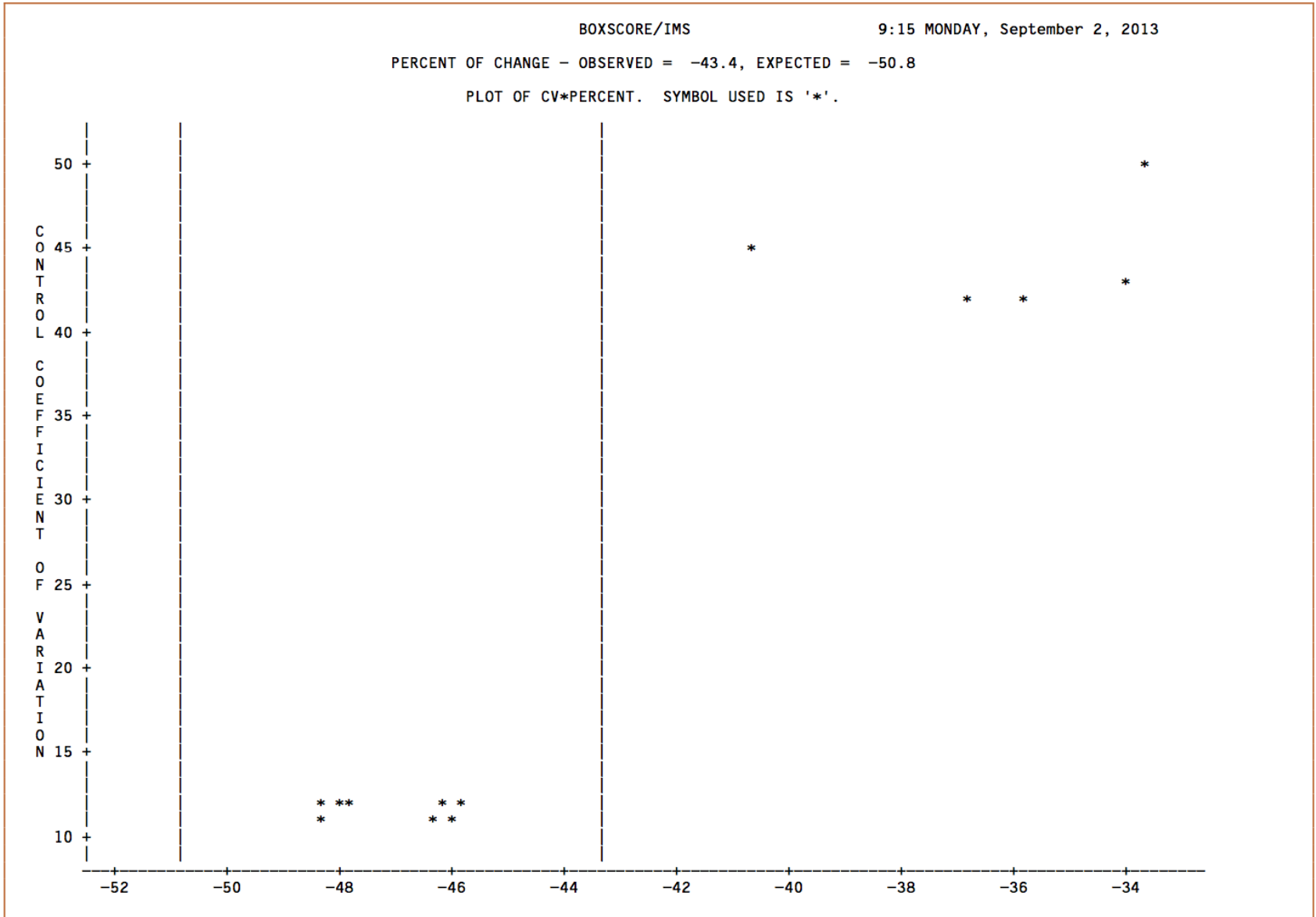
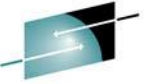


BOXSCORE/BATCH 10:16 Tuesday, December 4, 2012 9  
Percent of Change - Observed = 127.9, Expected = 132.6  
Plot of CV\*PERCENT. Symbol used is '\*'.  
Control  
Coefficient  
of  
Variation



# CPU per I/O

- First plot
  - BoxScore report showing a CPU upgrade that expected a CPU speed increase of 132.6%, but observed only a 127.9% increase – it was under-performing by about 5%.
- Second plot
  - BoxScore report showing a move from a z9 to a z114 that expected a drop in CPU speed of 50.8%, but saw only a drop of 43.4% - it was over-performing by about 7%. Each point represents one type of transaction of about 50,000 occurrences each.
  - What kind of normalization factor would work here?
    - None – some customers would be happy; some not so!





# Agenda

- CPU per I/O
- Why This Topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- z/OS Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

# Why This Topic?

- Recent Hardware Changes
  - IBM usually markets a new processor line every year, alternating the business class and enterprise class models
  - The average customer upgrades a processor every two to four years
  - Due to the amount of effort and cost, most customers wait for a CPC upgrade to also upgrade channels, network lines, coupling facilities, and memory, so many changes are being applied at one time
  - zPCR (WSC tool to help estimate capacity for an upgrade) is based on benchmarks that keep everything (CPU busy, channels utilization, memory usage, etc.) the same but for the CPU speed constant

# Why This Topic?

- Recent Customer Experiences
  - Many moves from z9 to z114 or from z10 to z196, have provided better savings (i.e. more capacity) than zPCR estimated
  - Many moves from z9 or z114 to zBC12 or from z10 or z196 to zEC12, have provided more capacity than zPCR estimated
  - The differences have been dramatic – up to 35% (even 40% in one case) different
  - Outsourcers are being hurt by a drop in revenue; nobody understands what's happening

# Agenda

- CPU per I/O
- Why This Topic?
  - Recent hardware changes
  - Recent customer experiences
- **Hardware Changes Affecting Variability**
- z/OS Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

# Hardware Changes Affecting Variability

- zIIPs/zAAPs (specialty processors or SPs)
  - Run at full speed even if running on knee-capped, or subcapacity CPCs (i.e. CP on zEC12 4xx is about 16% the speed of the 7xx, which is also the speed of the SP)
  - A job might run one day using CPs and SPs and the next day using only CPs; CPU time will differ; data sources will differ in their measurements (not all data sources track SP time correctly, if at all)
  - Normalization factor for SPs aren't perfect
  - Slight CPU overhead in switching to an SP, but could be reduced CPU if on subcapacity CPC, cost savings often seen in software and hardware pricing

# Hardware Changes Affecting Variability

- zEC12
  - Transactional execution exploited by Java 7 for z/OS and COBOL Compiler for z/OS V5.1 – result is decreased CPU times for Java users and decreased CPU time for programs recompiled under the new COBOL compiler
  - 2 GB Page Frames
    - Reduced CPU time for users of DB2 buffer pools and Java heap
  - Decimal floating point zoned conversion facility can reduce CPU time for jobs compiled with the new PL/I compiler
  - Look ahead instruction paths can reduce CPU time of a job

# Hardware Changes Affecting Variability

- General types of changes
  - Amount of cache in each level of memory, and the number of levels of memory, and the reference pattern of jobs (relative nest intensity) determine whether a specific job will run better or worse than other jobs; lots of variability here!
  - Location of instructions on chip can affect speed. As one example, the initial CMOS machines had some COBOL programs that took many times longer than expected; problem tracked down to programs using subscripts instead of indexes and the instructions of CVB and CVD had been moved off (or farther from) the CP that used those instructions.

# Hardware Changes Affecting Variability

- HiperDispatch
  - If turned on, CPU time of many jobs is generally reduced (from 0% to 10%)
  - This effect changes with each hardware release
  - Effect of poorly tuned LPARs (LP to CP ratio) is minimized with HD
- Coupling Facility (CF)
  - Speed of links affects CPU overhead of using CFs
  - Speed of coupling facility affects CPU overhead of data sharing jobs
  - Location of CF (internal vs external) – sometimes trade off between performance, cost, and reliability (single point of failure)



## Coupling Technology versus Host Processor Speed

Host effect with primary application involved in data sharing

Chart is based on 9 CF ops/Mi – may be scaled linearly for other rates

CF/Host	z10 BC	z10 EC	z114	z196	zBC12	zEC12
z10 BC ISC3	16%	18%	17%	21%	19%	24%
z10 BC 1x IFB	13%	14%	14%	17%	18%	19%
z10 BC 12x IFB	12%	13%	13%	15%	15%	17%
z10 BC ICB4	10%	11%	NA	NA	NA	NA
z10 EC ISC3	16%	17%	17%	21%	19%	24%
z10 EC 1x IFB	13%	14%	14%	17%	17%	19%
z10 EC 12x IFB	11%	12%	12%	14%	14%	16%
z10 EC ICB4	10%	10%	NA	NA	NA	NA
z114 ISC3	16%	18%	17%	21%	19%	24%
z114 1x IFB	13%	14%	14%	17%	17%	19%
z114 12x IFB	12%	13%	12%	15%	15%	17%
z114 12x IFB3	NA	NA	10%	12%	12%	13%
z196 ISC3	16%	17%	17%	21%	19%	24%
z196 1x IFB	13%	14%	13%	16%	16%	18%
z196 12x IFB	11%	12%	11%	14%	14%	15%
z196 12x IFB3	NA	NA	9%	11%	10%	12%
zBC12 ISC3	16%	17%	17%	21%	19%	24%
zBC12 1x IFB	14%	15%	14%	18%	17%	20%
zBC12 12x IFB	13%	13%	12%	15%	14%	17%
zBC12 12x IFB3	NA	NA	10%	11%	11%	12%
zEC12 ISC3	16%	17%	17%	21%	19%	24%
zEC12 1x IFB	13%	13%	13%	16%	16%	18%
zEC12 12x IFB	11%	11%	11%	13%	13%	15%
zEC12 12x IFB3	NA	NA	9%	10%	10%	11%

With z/OS 1.2 and above, synch-> asynch conversion caps values in the table at about 18%  
 IC links scale with the speed of the host technology and provide an 8% effect in each case

# Hardware Changes Affecting Variability

- zBC12 and zEC12 GA2 (September 2013)
  - New Integrated Firmware Processor (IFP) processor
    - Used for native PCIe functions such as zEDC Express and 10GbE RoCE Express
  - Newer faster FICON Express8 and 8S
    - Improved channel speed usually reduces the CPU time of a job
  - OSA 10 GbE, 1000BASE-T
    - Improved network speed usually reduces the job CPU time
  - zEnterprise Enhanced Data Compression (zEDC)
    - Compression, even H/W compression, can increase OR decrease the CPU time used by a job

# Hardware Changes Affecting Variability

- zBC12 and zEC12 GA2 (September 2013)
  - Unified Resource Manager
    - Additional traffic from zBX or IDAA could increase the job CPU time
  - More memory
    - More memory can sometimes decrease some jobs' CPU time (fewer I/Os because of buffering and elapsed time is decreased), but it can also increase some jobs' CPU time (e.g. SORT using in-storage sort instead of using I/Os)
  - Flash Express
    - This acts like a faster paging device and usually results in less CPU time per job

# Agenda

- CPU per I/O
- Why This Topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- z/OS Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

# z/OS Changes Affecting Variability

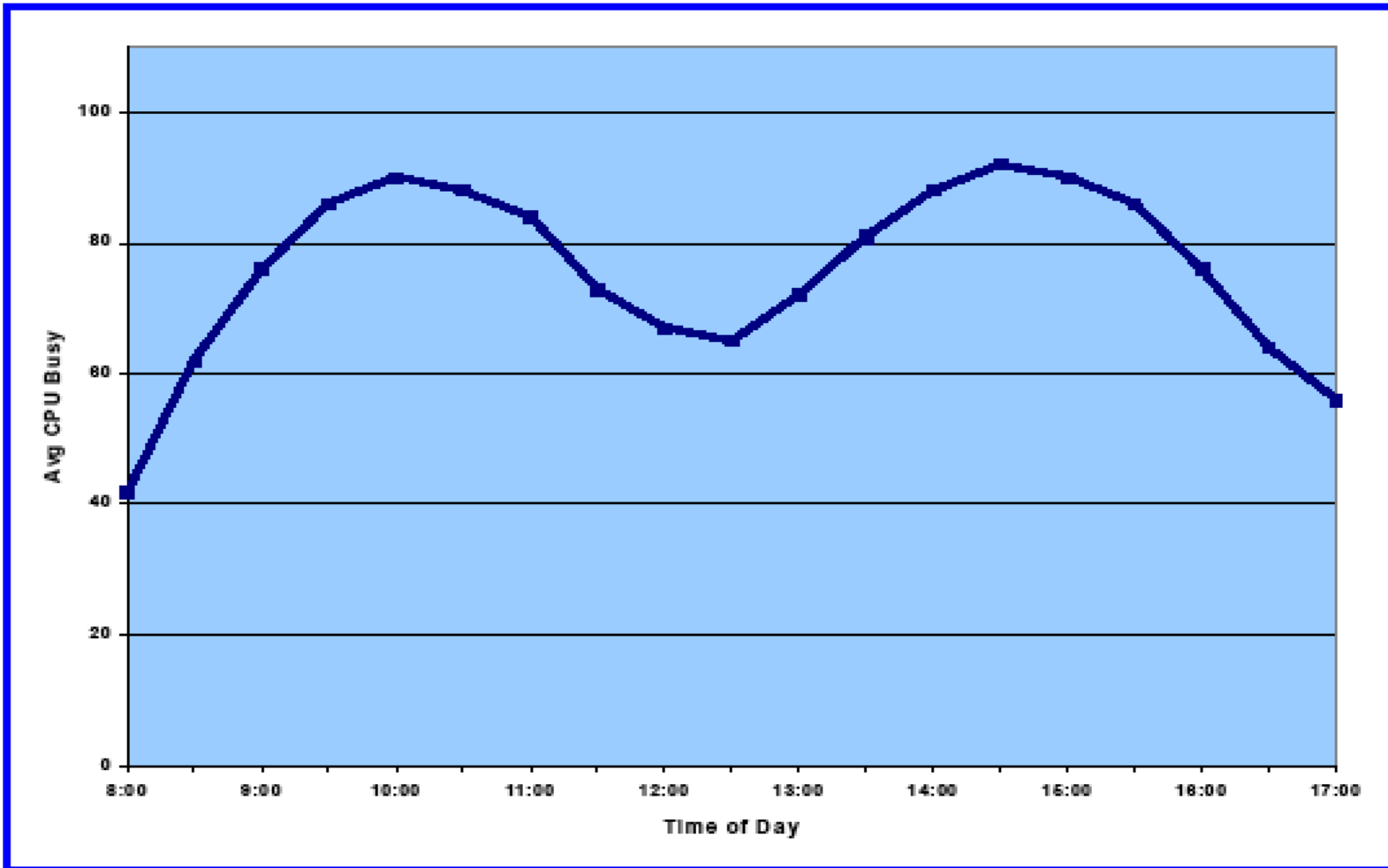
- Operating system levels affect CPU time (usually each release reduces it by a certain percent).
- Constrained resources increases CPU time in jobs (e.g. slow DASD or constrained storage).
- Maintenance may introduce performance APARs that will affect the amount of CPU time consumed, especially if these are new function APARs.
- Maintenance of vendor products may introduce changes in CPU times.

# Agenda

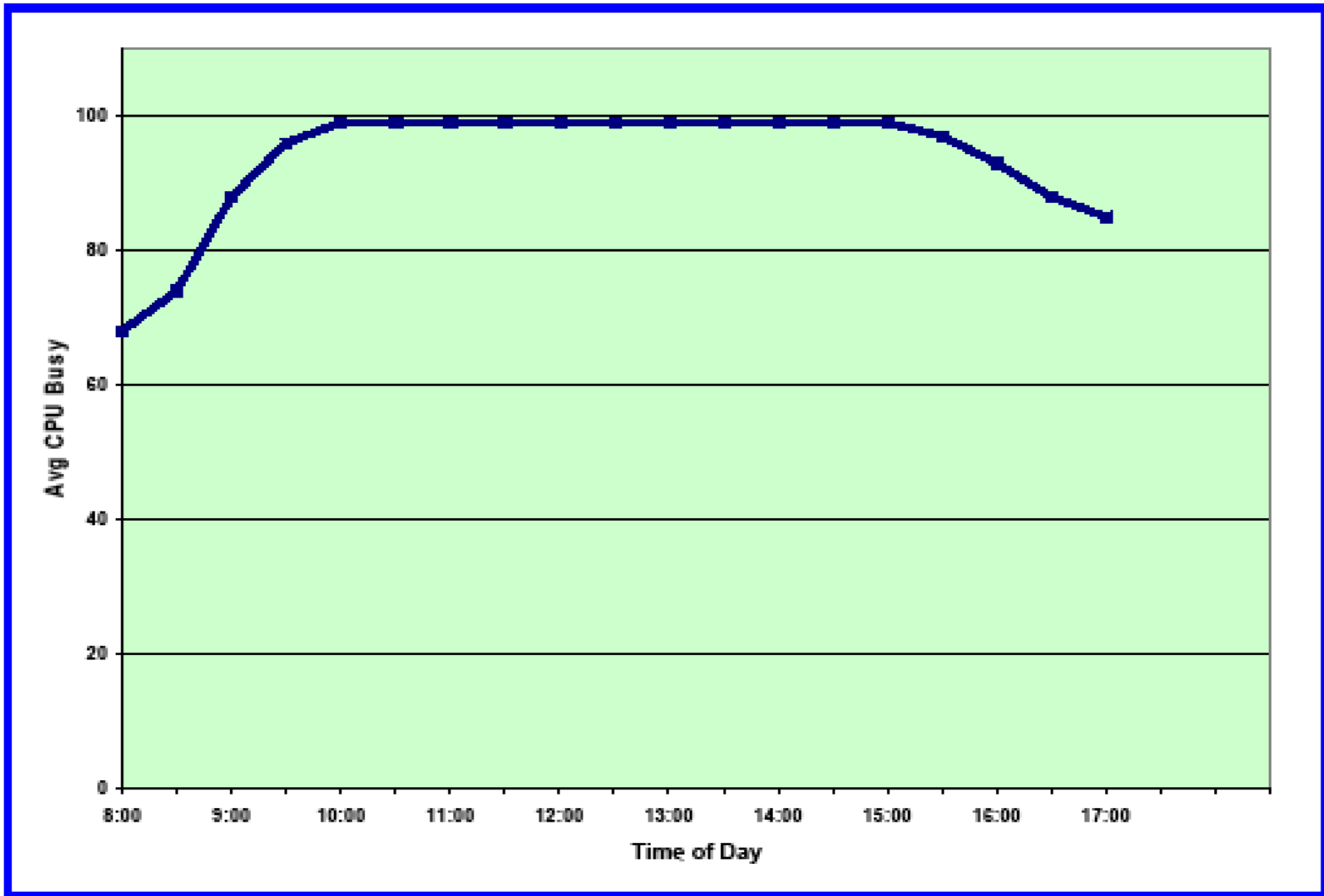
- CPU per I/O
- Why This Topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- z/OS Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

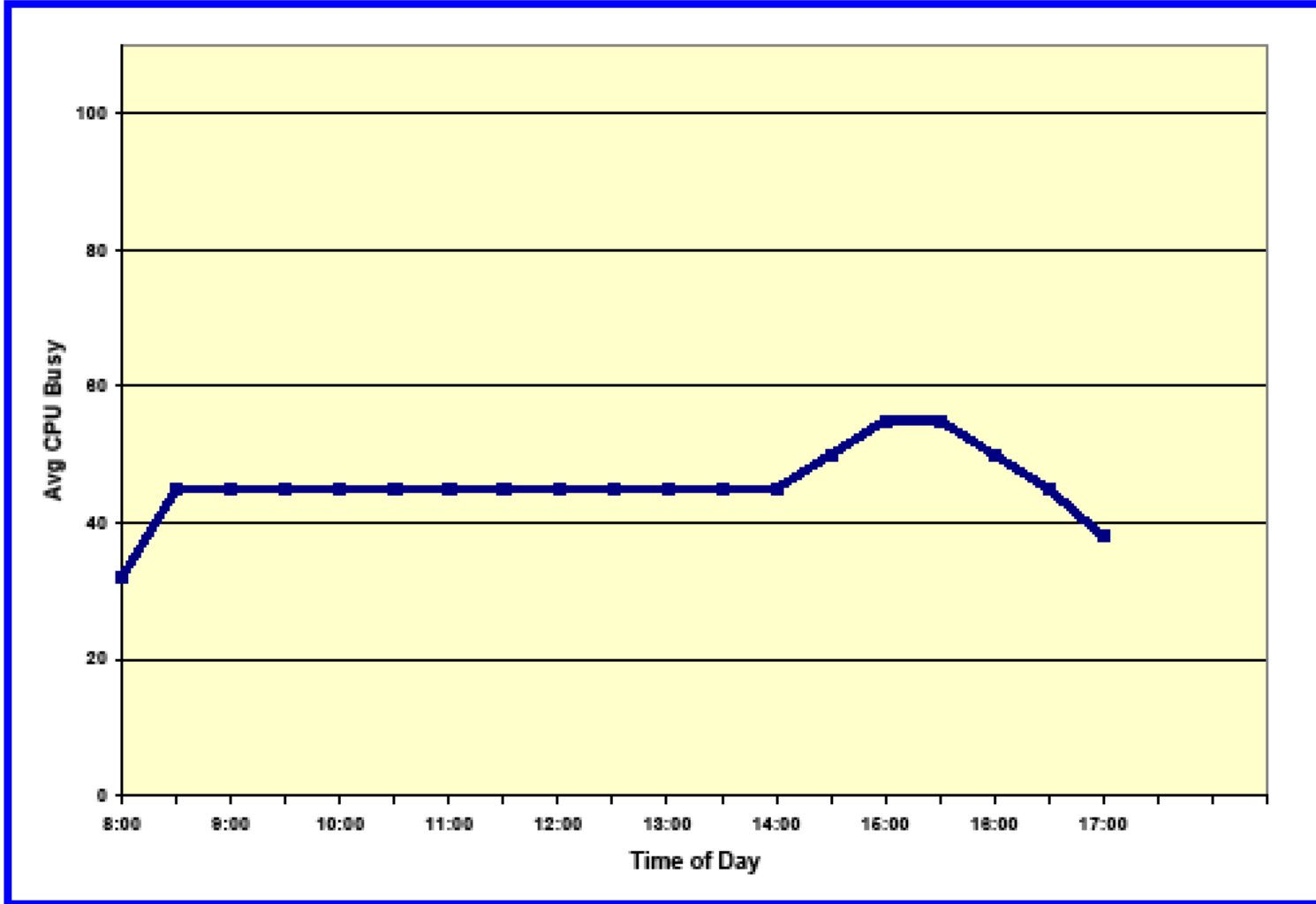
# Environment Changes – CPU Busy

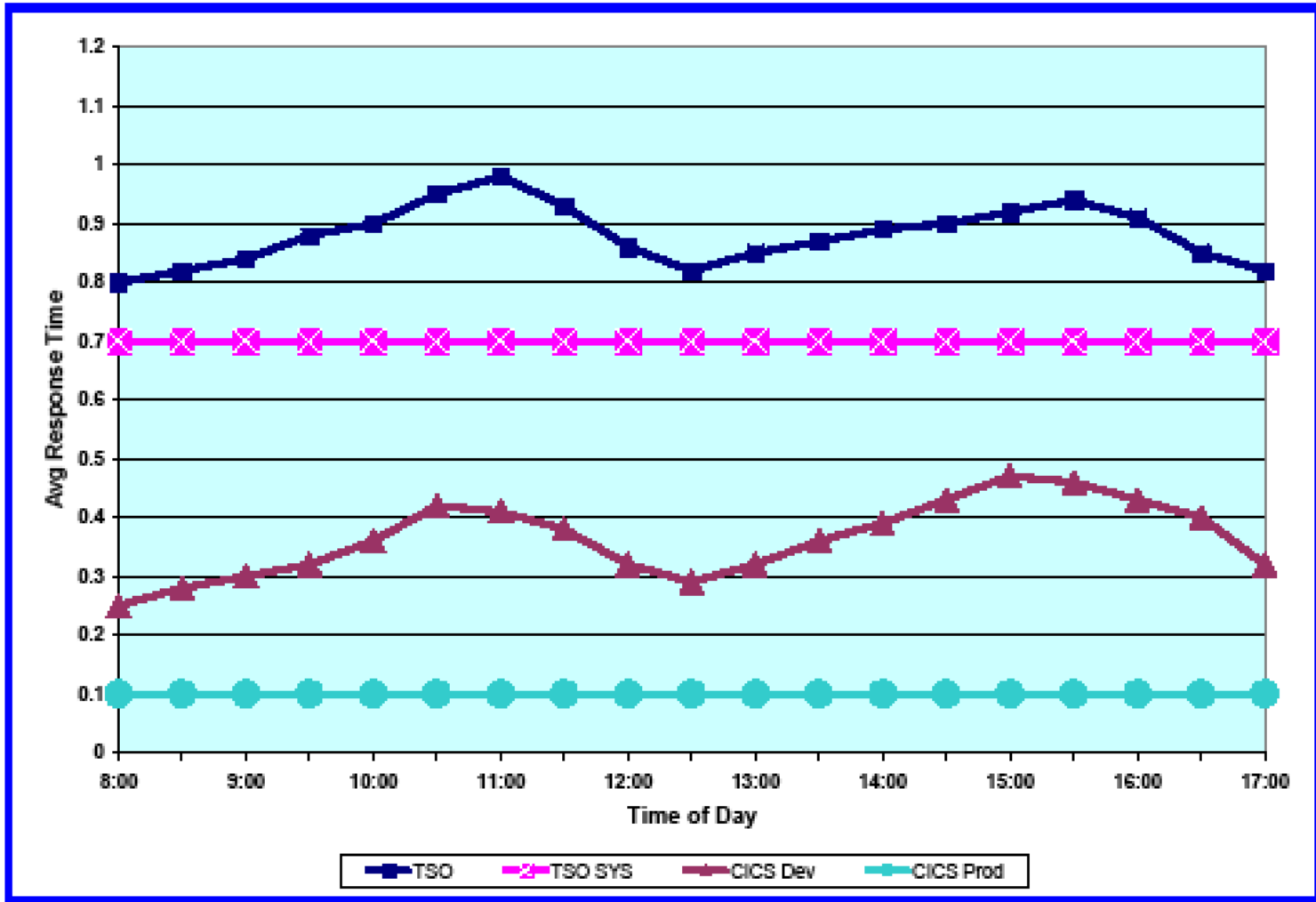
- CPU time of jobs and transactions are affected by CPU utilization. An increase of 10% in physical CPU busy can increase job CPU time by 3-5%.
- IBM's LSPRs are determined for batch and online at 90% busy and mixed workloads at 99% busy. If your system is at 70% busy, your jobs could take up to 20% less CPU time than you expect from zPCR.
- Right after an upgrade, many sites run at a lower CPU utilization.
- Compare the following graphs of the traditional “Sofía Vergara” CPU busy, the more current “Hulk Hogan” chart, and a typical latent demand chart. (All from *Tuning Letter 2003 No. 6*)

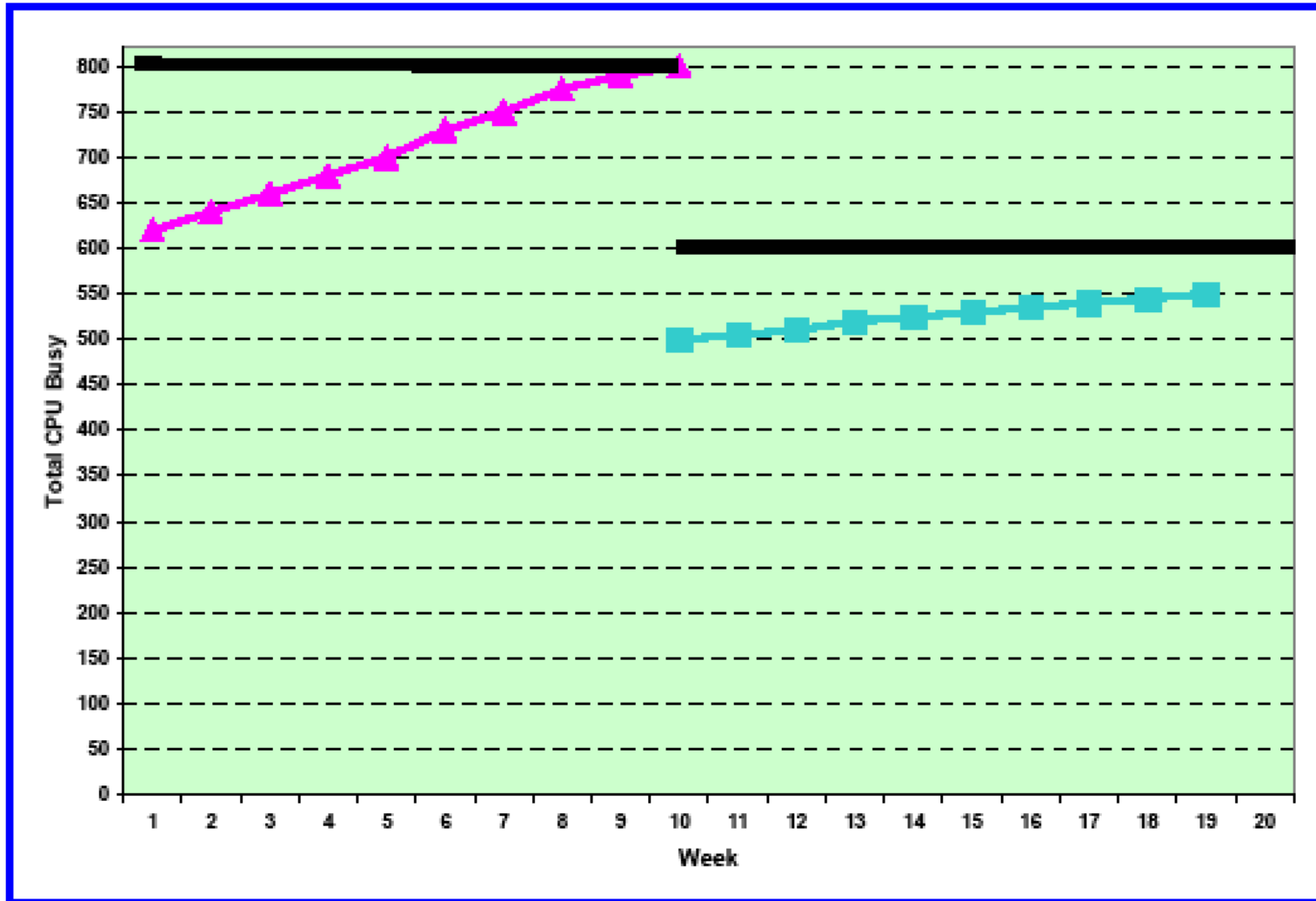












# Environment Changes – Other Factors

- Increase of LPAR weight can reduce job CPU time; decrease of weight can increase job CPU time
- Job run in larger LPAR will take less CPU time than when run in small LPAR
- Changes in workloads in other LPARs, number of LPARs, number of LPs in all LPARs can also change job's CPU time

# Agenda

- CPU per I/O
- Why This Topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- z/OS Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

# Other Changes Affecting Variability

- Application tuning can change CPU times.
- DASD tuning can change CPU times. (E.g. implementing system determined blocksizes)
- Change in database size, especially indexed VSAM files, can change job CPU times
- APAR maintenance – many APARs provide performance improvements
- Compiler levels and options
- Number of synchronous vs asynchronous requests
- Number of interrupts on the system
- Room temperature
- “I didn’t change anything!” changes

# Agenda

- CPU per I/O
- Why This Topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- z/OS Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What to Do?



# Measurements

- How do you determine what to expect?
  - Sites SHOULD be using zPCR to determine expectations.
  - Unfortunately many sites don't. Even so, which measurements do you use?

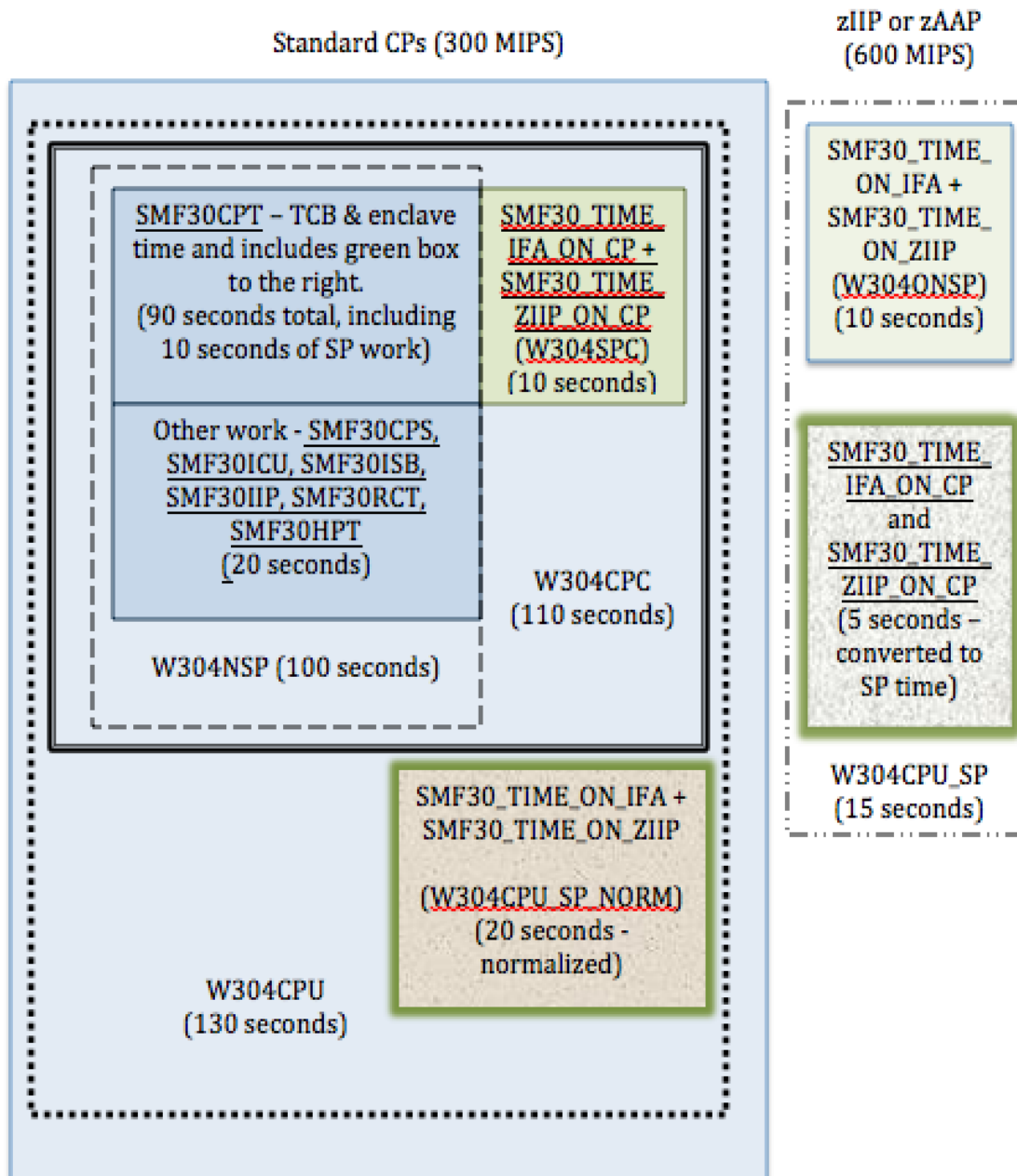
Item	z10-BC-U03	zBC12-005	Delta	% Delta	ITRR
Speed of one CPU (logical):					
Expected SU/second	18285.71	14134.27	-4151.44	-22.7%	0.77
Expected avg MIPS/CPU	346.8	277.1	-69.8	-20.1%	0.80
Expected max MIPS/CPU	384.1	291.1	-93.1	-24.2%	0.76
Expected min MIPS/CPU	278.3	246.3	-32.0	-11.5%	0.88
Observed MIPS/CPU	346.8	308.5	-38.4	-11.1%	0.89
Change from expected avg			31.4	11.3%	1.11
Weighted MIPS/CPU	346.8	327.1	-19.7	-5.7%	0.94
Change from weighted avg			50.0	18.1%	1.18

# Measurements

- References
  - IBM SRM constants page: <https://www-304.ibm.com/servers/resourceLink/lib03060.nsf/pages/srmindex?OpenDocument&pathID=>
  - IBM LSPR page: <https://www-304.ibm.com/servers/resourceLink/lib03060.nsf/pages/sprindex?OpenDocument>
  - *Cheryl Watson's System z CPU Chart*

# Measurements

- The fields you use for measurement could be more or less stable.
  - One example is from our CPU measurements presentation (next slide). See [www.watsonwalker.com/PR131204.pdf](http://www.watsonwalker.com/PR131204.pdf).



# Agenda

- CPU per I/O
- Why This Topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- z/OS Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

# What To Do?

- Measure, measure, measure!
- Run daily reports with CPU usage by CPC, LPAR, workload to understand variance by hour, by day, by week, by month, by year, by CPU utilization. Understand what's "normal."
- Identify major changes.
- Benchmark (I prefer CPU per I/O method to select jobs) before and after any change.
- Use zPCR to estimate configuration changes; but realize that not all variables are known by zPCR.
- Use zSoftCap to estimate changes in software or subsystem releases.
- Cap your LPARs after migrating to a new processor so that you aren't running at a low CPU utilization.
- Consider new options for chargeback, such as tiering.

# New in z/OS 2.1

- SMF Type 30, Counter Section
  - Activated when SMFCOUNT is specified in SMFPRMxx (or set with SETSMF command) and Hardware Instrumentation Services (HIS) is running
  - Records number of instructions executed on:
    - CP as TCB (non-enclave)
    - CP as SRB (non-enclave)
    - CP as preemptable or client SRB (non-enclave)
    - zIIP/zAAP (non-enclave)
    - CP but eligible for zIIP/zAAP (non-enclave)
    - CP as independent enclave
    - zIIP/zAAP as independent enclave
    - CP but eligible for zIIP/zAAP as independent enclave
    - CP as dependent enclave
    - zIIP/zAAP as dependent enclave
    - CP but eligible for zIIP/zAAP as dependent enclave
  - IBM expected these counts to be relatively repeatable, but they don't seem to be. They're looking for more sample SMF data to confirm.

## See John Burg's Session [15705](#) (Thu, 11:15, CPU MF Update)

### Looking for zOS 2.1 Instruction Count “Volunteers” to send SMF data

- **We want to determine Instruction Vs CPU time consistency at various utilizations**

#### Looking for “Volunteers”

- SMF data from 3 days from 2 different “like” weeks (e.g. Mon-Wed)
  - 24 hours/day, SMF 30s, 70s, 72s, 113s per LPAR
- z/OS 2.1 with “SMF30COUNT” enabled and CPU MF Counters enabled
- Production partitions only
  - Preferred Customer Profile:
    - Websphere, or CICS running with Transaction goals, or large DDF workload, or small utility type Batch jobs that run consistently thru day

If interested send note to [jpburg@us.ibm.com](mailto:jpburg@us.ibm.com),

No deliverable will be returned

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



# All Watson & Walker Sessions

- Thank you for coming!
- If you liked this session, please check out others from Cheryl and Frank:
  - [15602](#) – Wed, 1:30 pm – Frank – *The Skinny on Coupling Facility Interrupts*
  - [16251](#) – Thu, 3 pm – Cheryl & Frank – *Hot Tips From Cheryl and Frank*
  - [15567](#) – Fri, 10 am – Cheryl & Frank – *Exploiting z/OS Tales From the MVS Survey*
- If you like SMF data, please see our new *SMF Reference Summary* at [www.watsonwalker.com/references.html](http://www.watsonwalker.com/references.html)