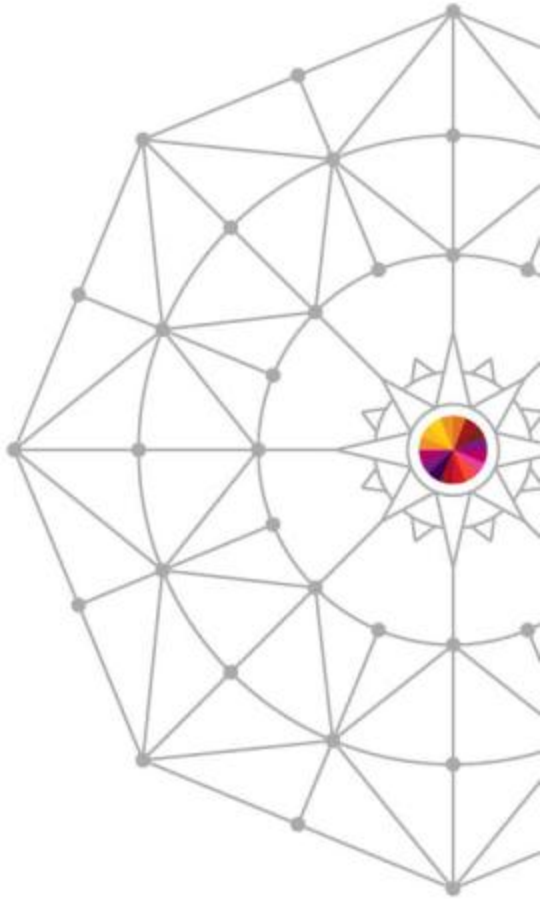
#SHAREorg

SHARE

Why is the CPU Time for a Job so Variable?

Cheryl Watson
Watson & Walker, Inc.
www.watsonwalker.com
technical@watsonwalker.com

March 12, 2014
Session 15274



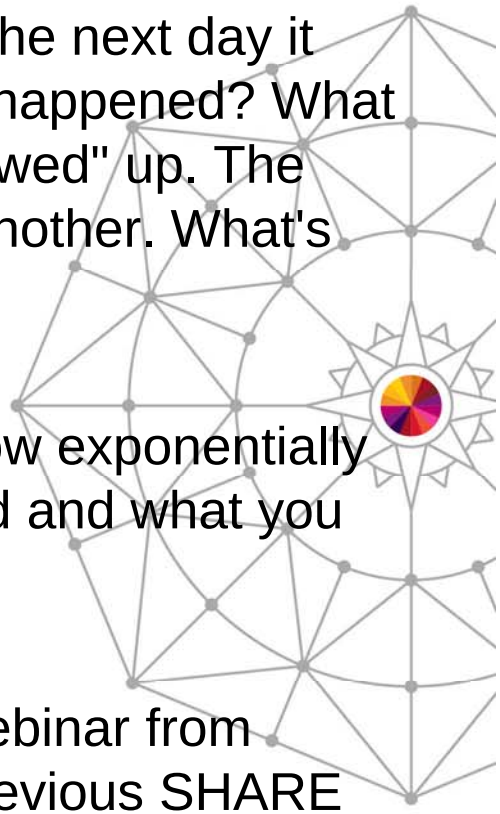
SHARE in Anaheim

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, who has been watching this problem grow exponentially since 1965, has 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.



Agenda

- Comparison of 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?

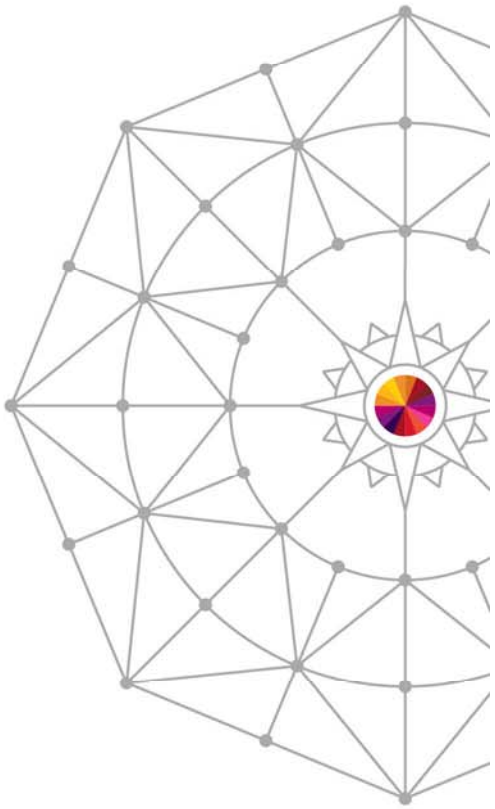
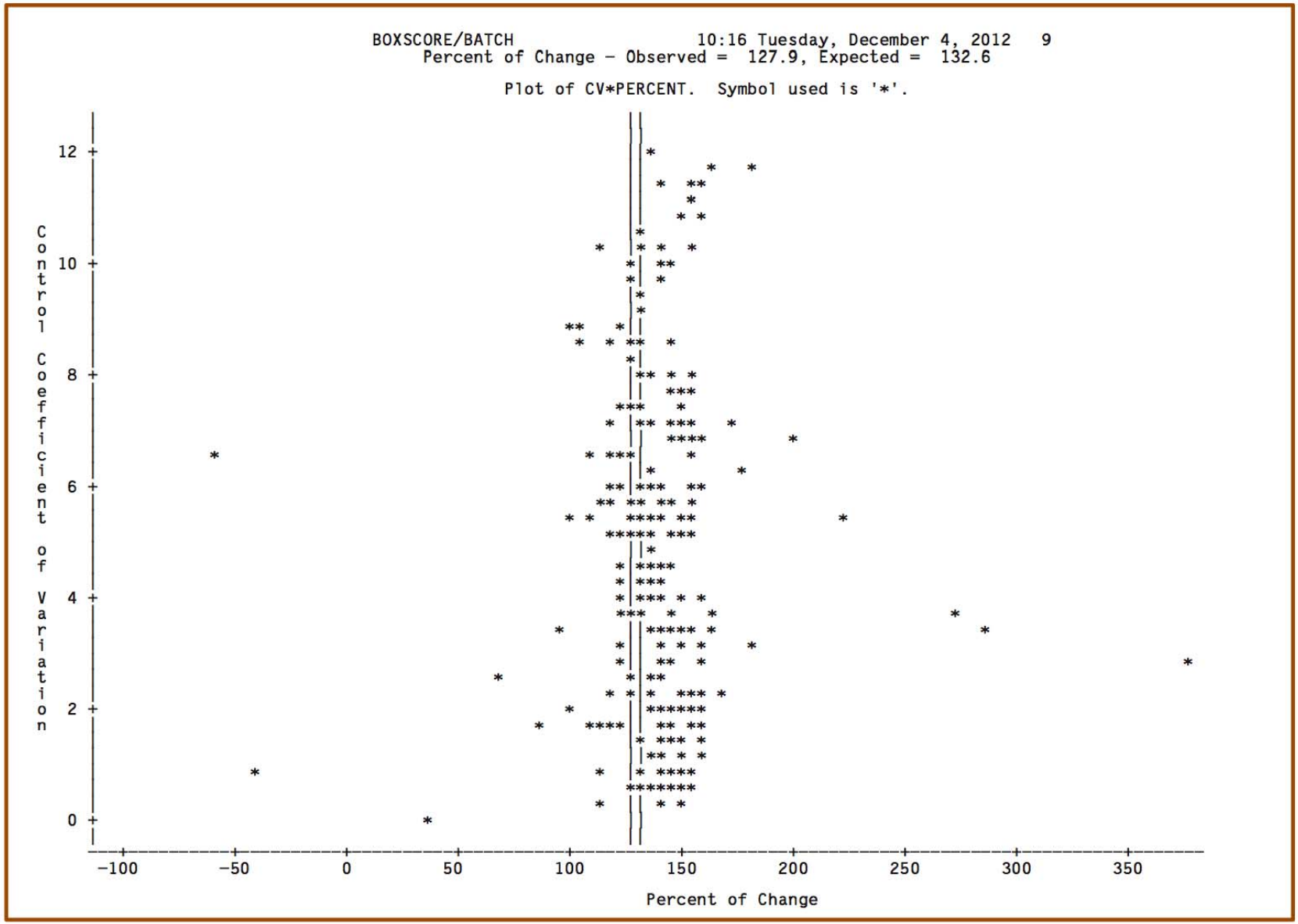


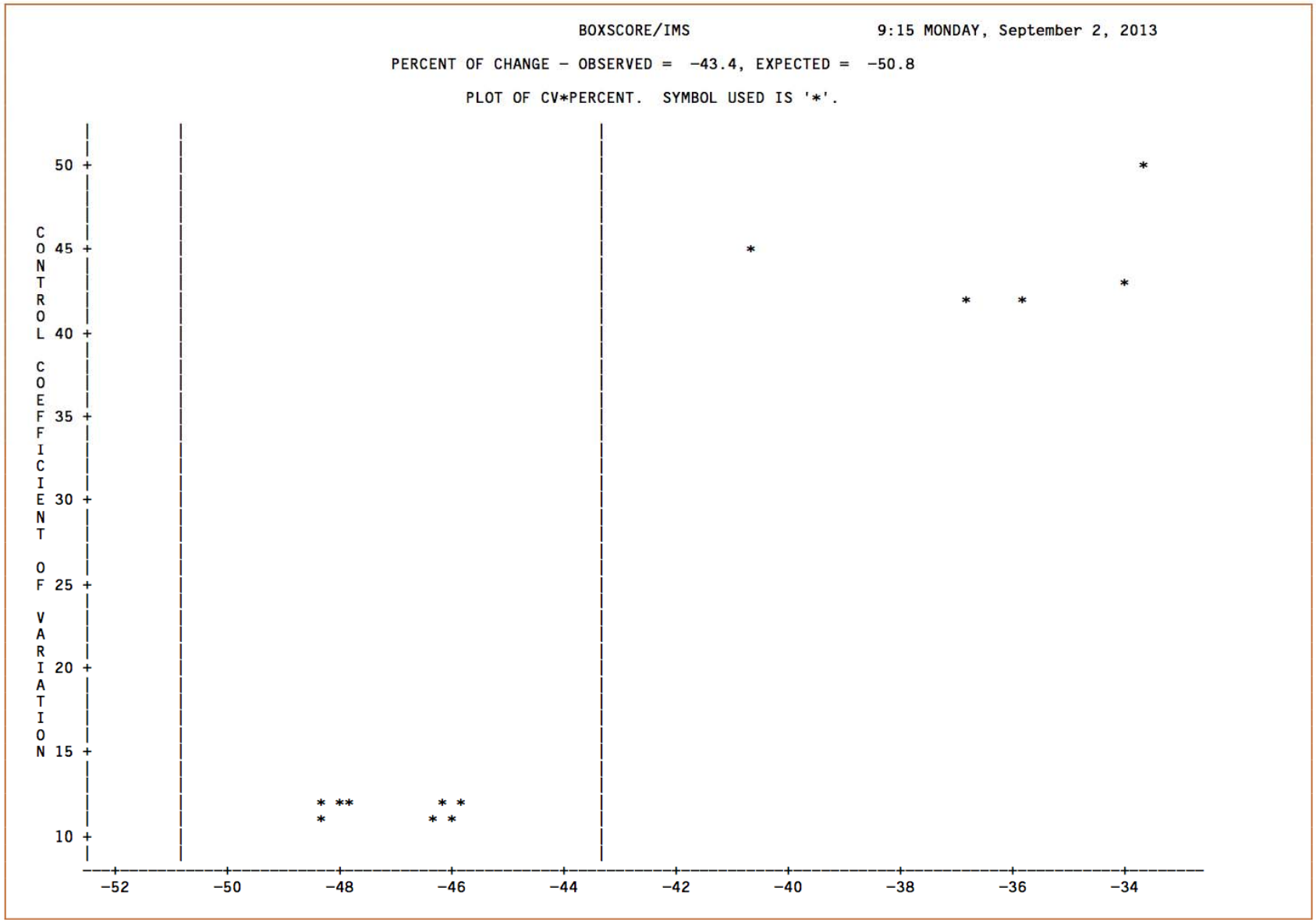
Figure 10 – BoxScore Chart of CPU Changes



Plot Reports Based on CPU per I/O

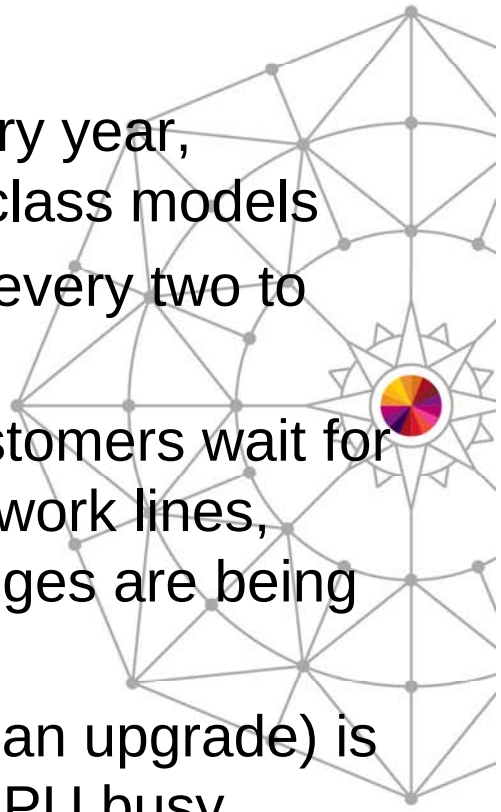
- First plot
  - BoxScore report showing a CPU upgrade that expected a CPU speed increase of 132.6%, but observed only a 127.9% increase – it was under-performing by about 5%.
- Second plot
  - BoxScore report showing a move from a z9 to a z114 that expected a drop in CPU speed of 50.8%, but saw only 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?

Figure 1 – BoxScore Plot of Percent of Change



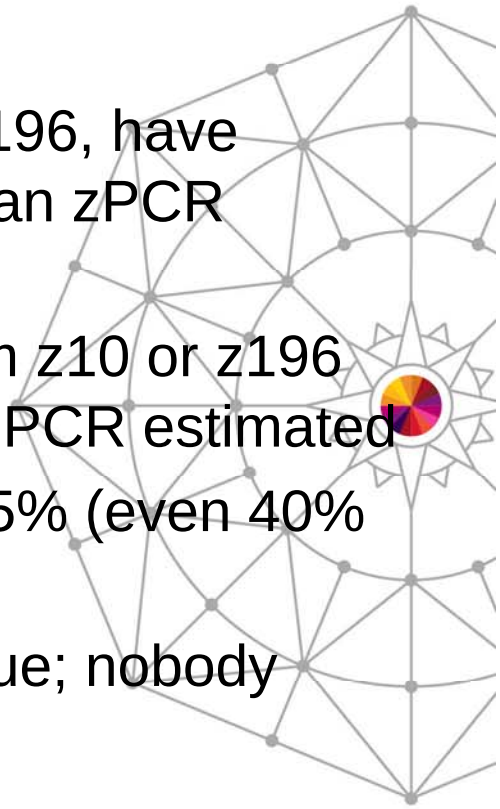
Why This Topic?

- 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, memory, etc.) but 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 once case) different
  - Outsourcers are being hurt by a drop in revenue; nobody understands what's happening



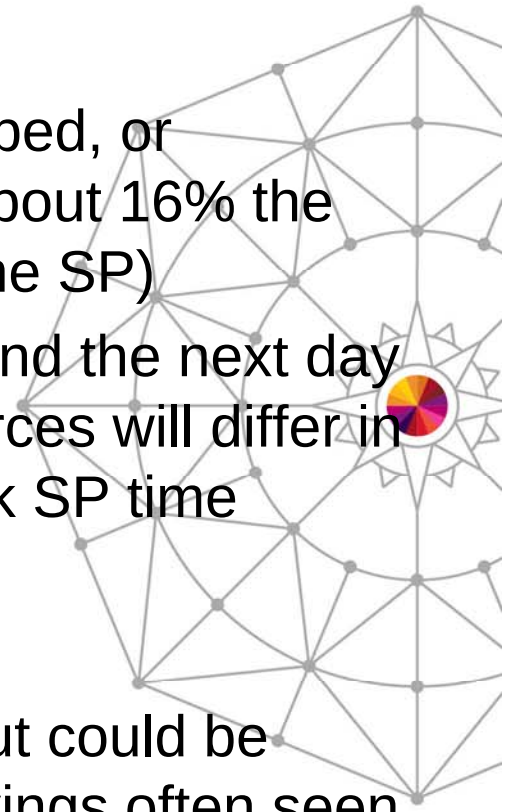
Newly Added Session

New!

- We added a vendor session (not in printed schedule) for Thursday at 1:30 pm that contains related information on outsourcing issues (plus other tips from me and Frank Kyne) – see Session 15397

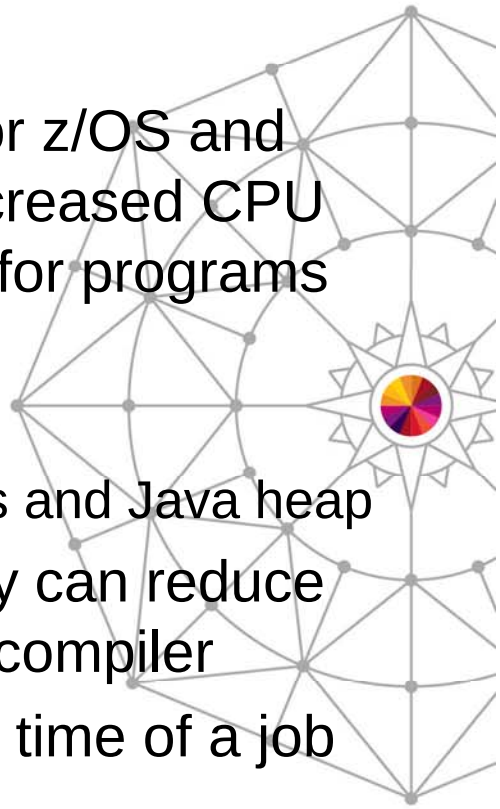
Hardware Changes Affecting Variability

- 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 uses and decreased CPU time for programs recompiled under new 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 (e.g. 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
  - This effect changes with each hardware release
  - Effect of poorly tuned LPARs (LP to CP ratio) is minimized with HD
- Coupling Facility
  - Speed of links affects CPU overhead of using coupling facility
  - Speed of coupling facility affects CPU overhead of data sharing jobs

See Gary King's session 13093 (SHARE 2013 SF)

Coupling Technology versus Host Processor Speed

Host effect with primary application involved in data sharing

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

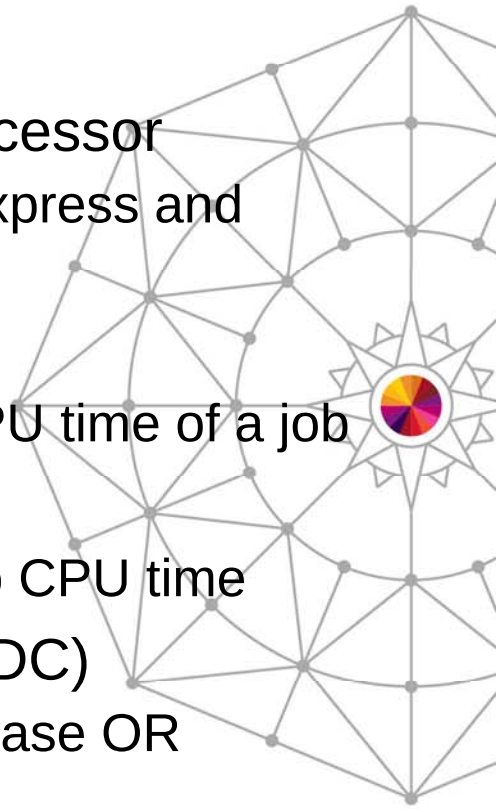
CF\Host	z10 BC	z10 EC	z114	z196	zEnterprise EC12
z10 BC ISC3	16%	18%	17%	21%	24%
z10 BC 1x IFB	13%	14%	14%	17%	19%
z10 BC 12x IFB	12%	13%	13%	15%	17%
z10 BC ICB4	10%	11%	NA	NA	NA
z10 EC ISC3	16%	17%	17%	21%	24%
z10 EC 1x IFB	13%	14%	14%	17%	19%
z10 EC 12x IFB	11%	12%	12%	14%	16%
z10 EC ICB4	10%	10%	NA	NA	NA
z114 ISC3	16%	18%	17%	21%	24%
z114 1x IFB	13%	14%	14%	17%	19%
z114 12x IFB	12%	13%	12%	15%	17%
z114 12x IFB3	NA	NA	10%	12%	13%
z196 ISC3	16%	17%	17%	21%	24%
z196 1x IFB	13%	14%	13%	16%	18%
z196 12x IFB	11%	12%	11%	14%	15%
z196 12x IFB3	NA	NA	9%	11%	12%
zEnterprise EC12 ISC3	16%	17%	17%	21%	24%
zEnterprise EC12 1x IFB	13%	13%	13%	16%	18%
zEnterprise EC12 12x IFB	11%	11%	11%	13%	15%
zEnterprise EC12 12x IFB3	9%	9%	9%	10%	11%

With z/OS 1.2 and above, synch->asynch conversion caps values in table at about 18%

IC links scale with speed of host technology and would provide an 8% effect in each case

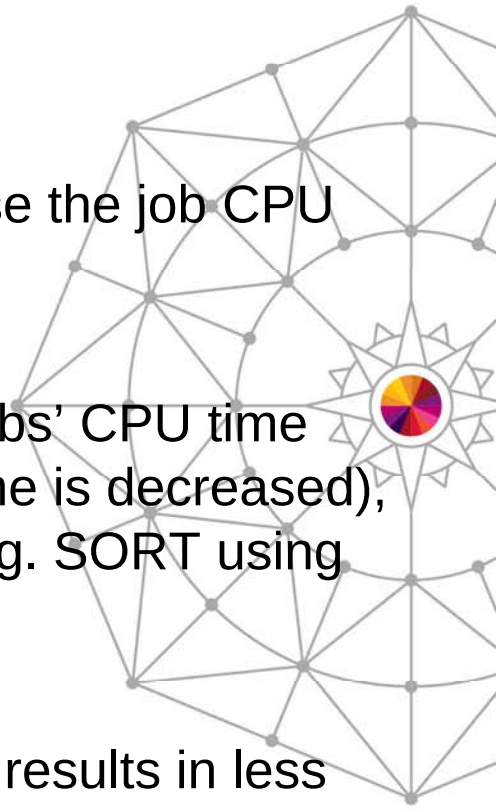
Hardware Changes Affecting Variability

- 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, 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



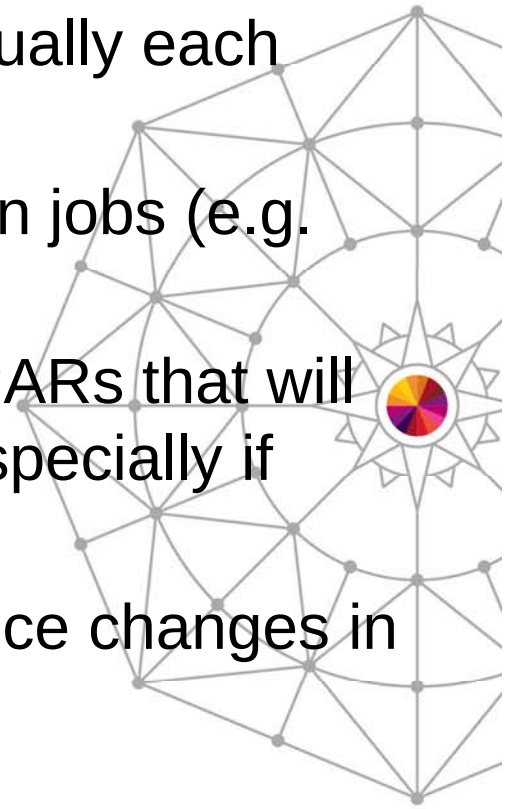
Hardware Changes Affecting Variability

- 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



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.



Environment Changes – CPU Busy

- CPU time of jobs and transactions are affected by CPU utilization. A increase of 10% in 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 "Dolly Parton" CPU busy, the more current "Hulk Hogan" chart, and a typical latent demand chart. (All from *Tuning Letter 2003 No. 6*)

Figure 4 - Average CPU Busy

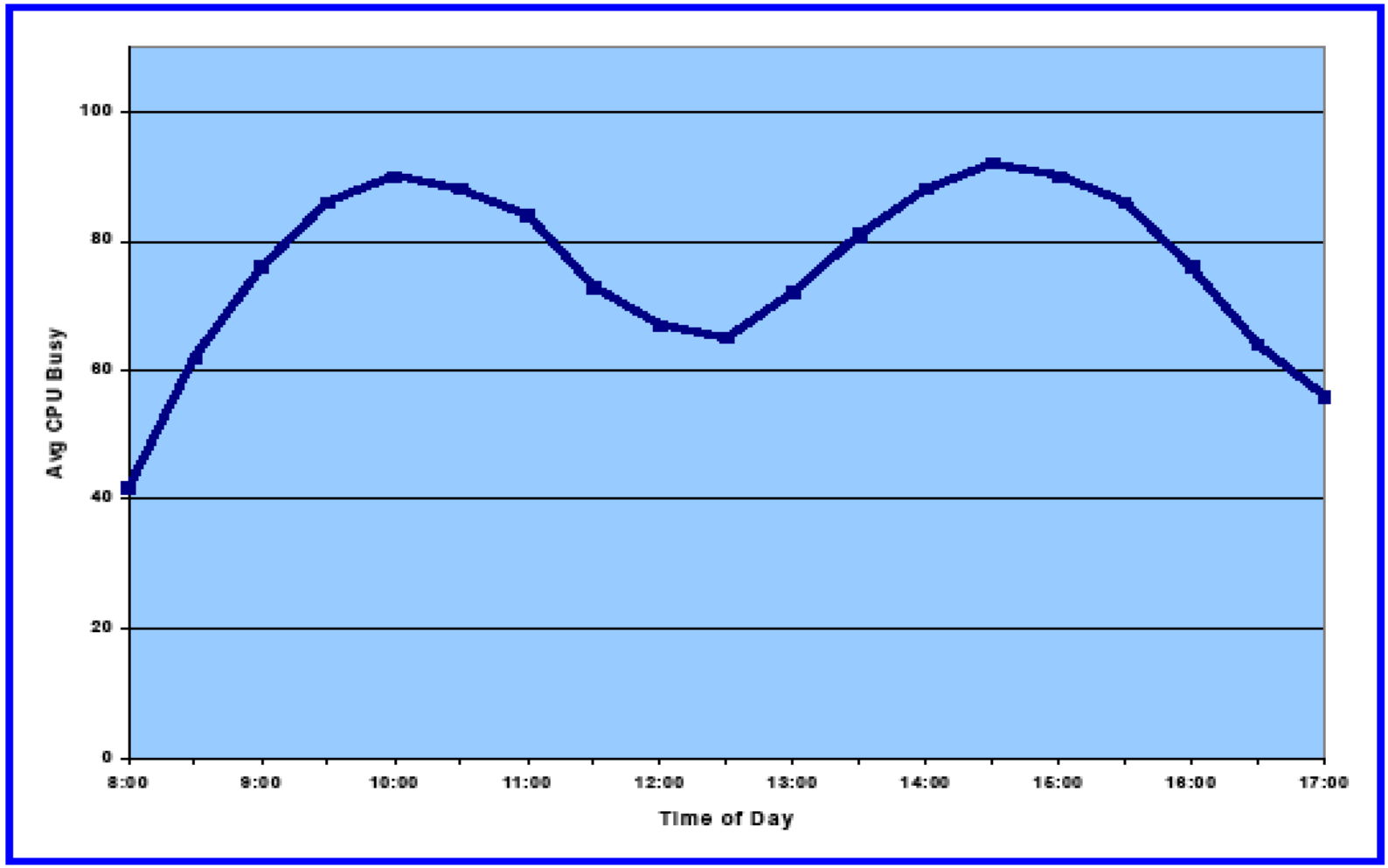


Figure 5 - Average CPU Busy - Saturated

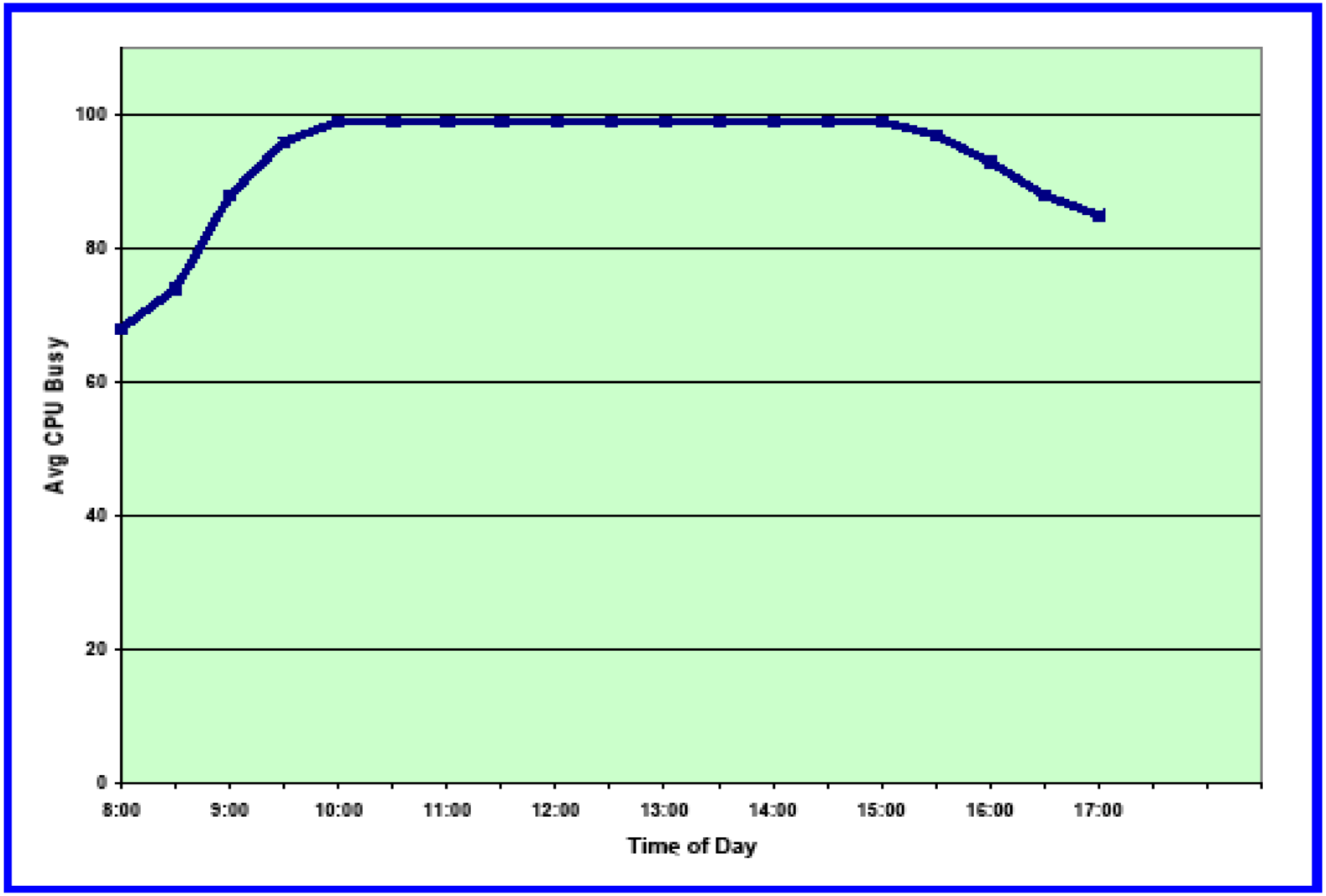


Figure 6 - Average CPU Busy - LPAR

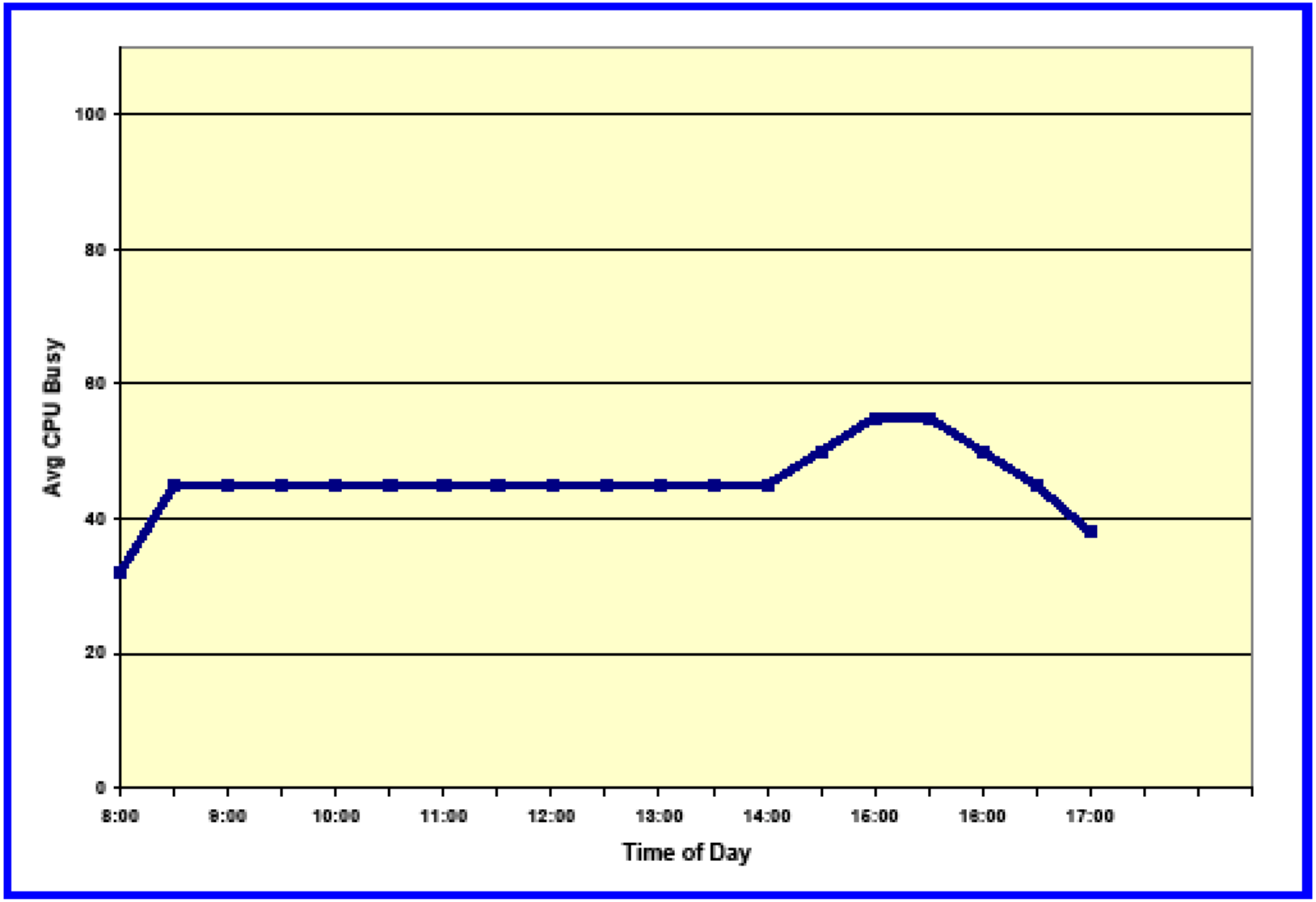


Figure 12- Average Response Times

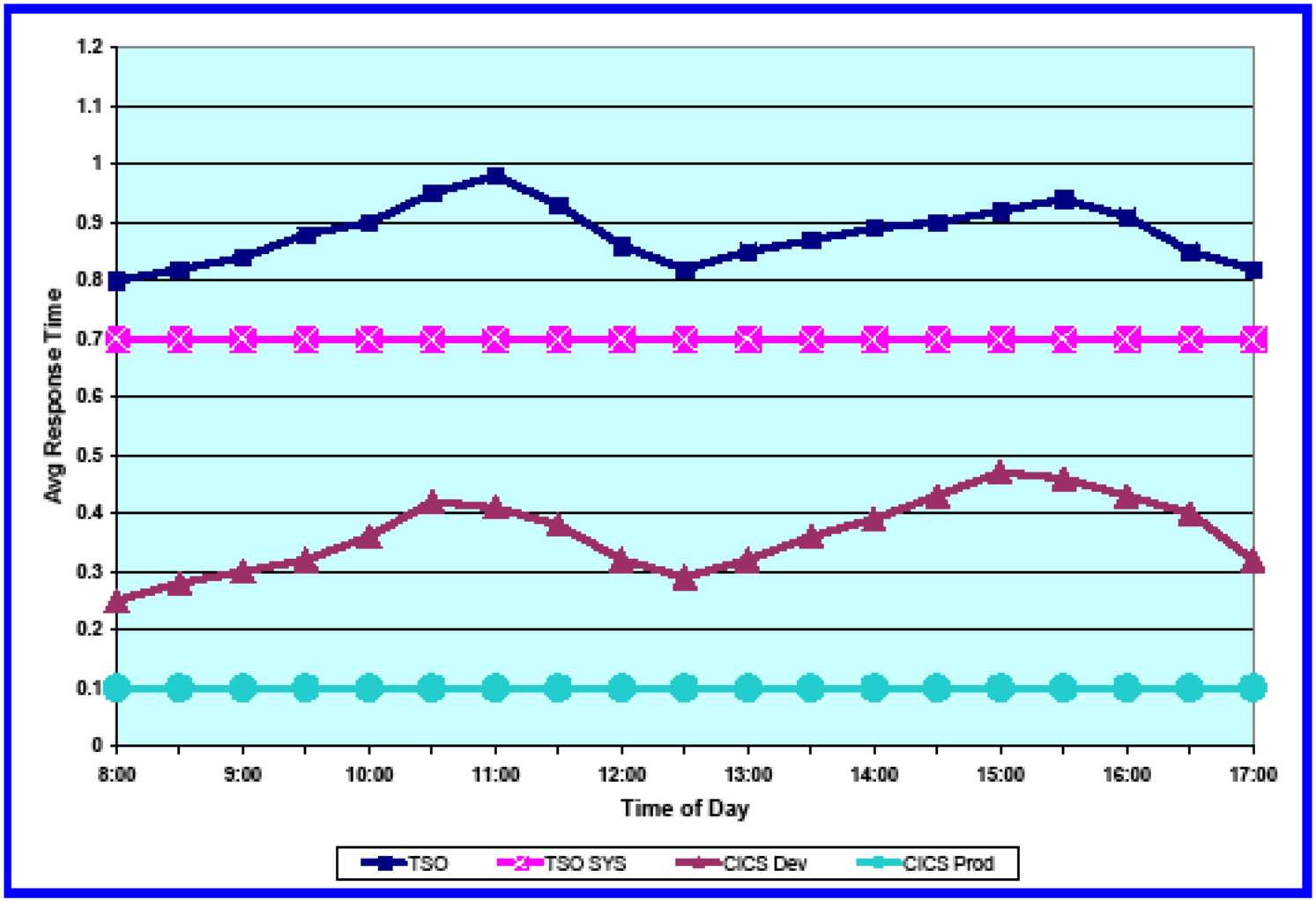
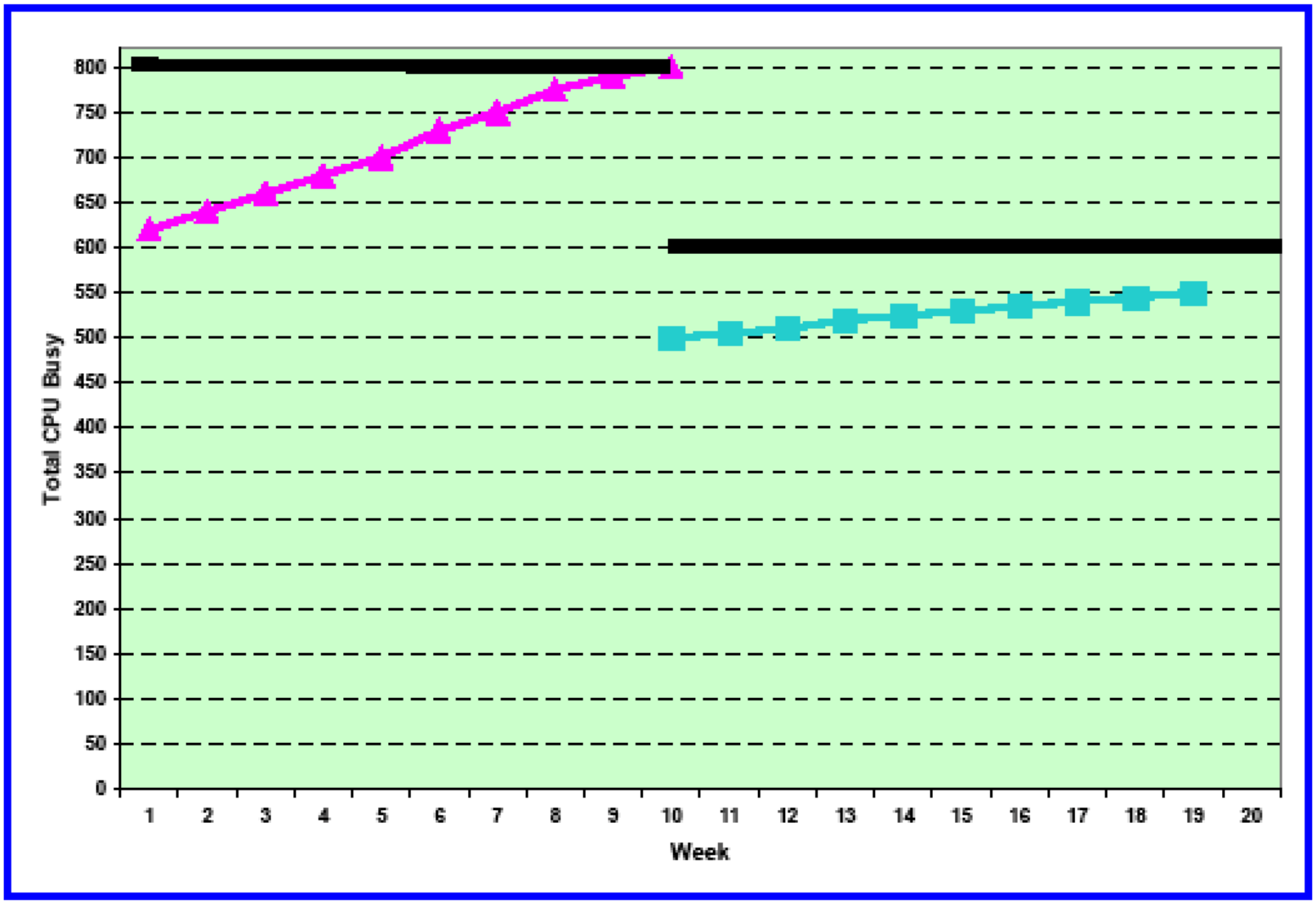
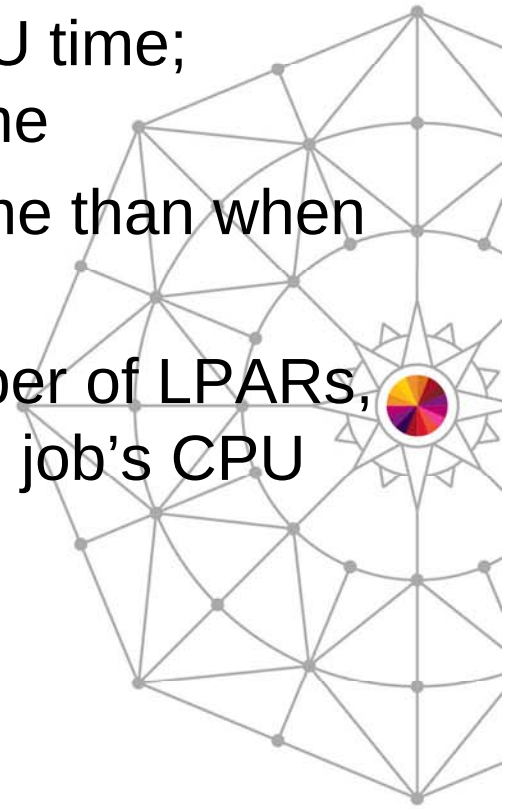


Figure 15 - Weekly Total CPU Busy - Changing Processors



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



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.

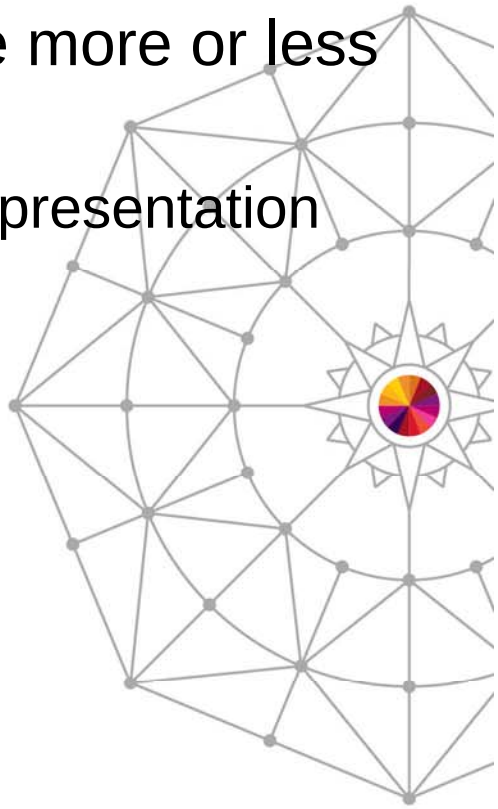
Measurements

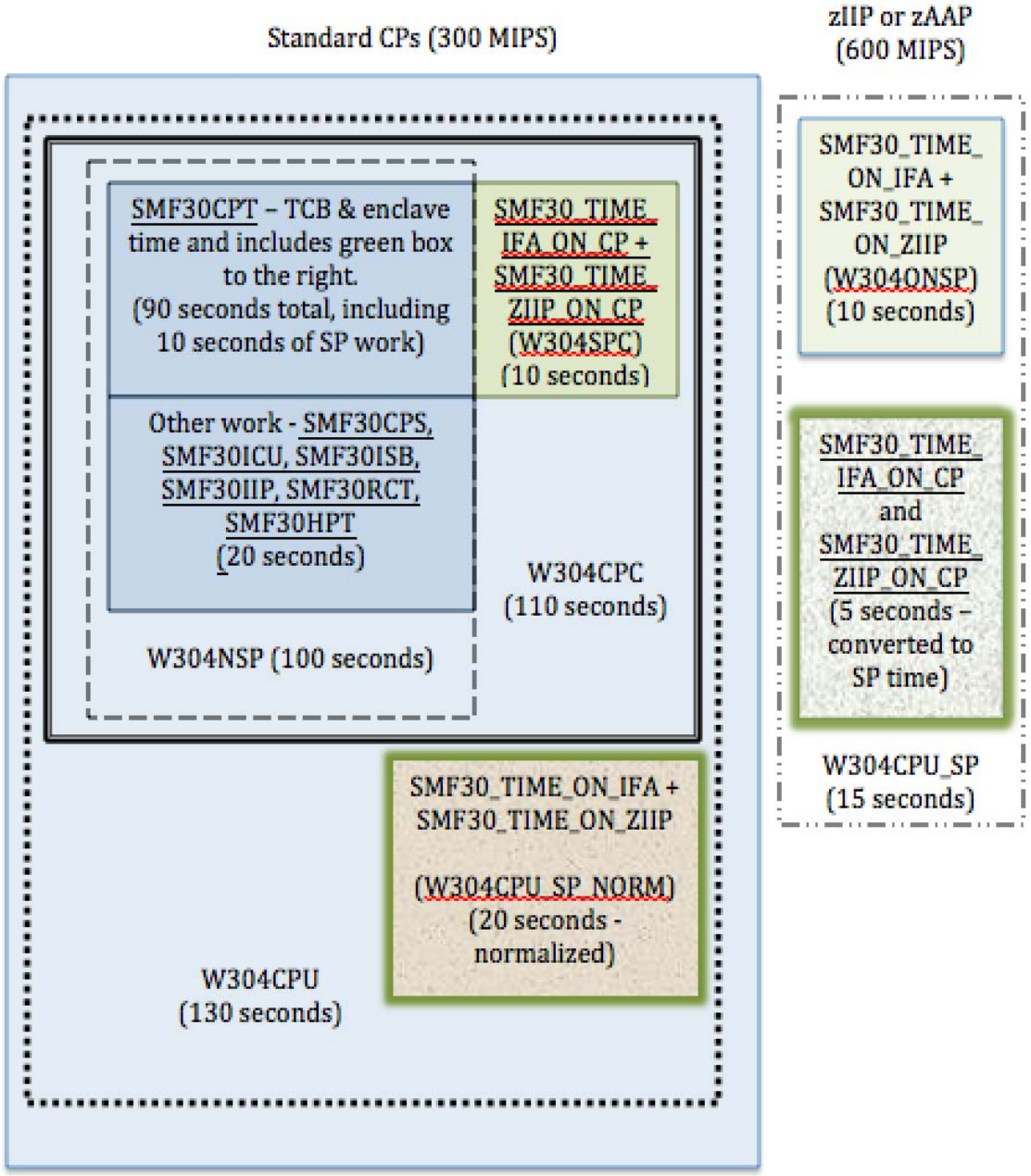
- How do you determine what to expect?
  - Sites SHOULD be using zPCR to determine expectations.
  - Unfortunately many sites don't. Even so, which measurements do you use?

Item	z10-BC-U03	zBC12-O05	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

- The fields you use for measurement could be more or less stable.
  - One example is from my CPU measurements presentation (next slide).





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. 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.
- Consider new options for chargeback, such as next slide.

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 enabled for the Basic Counter Set
  - 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