Frank Kyne
Editor and Technical Consultant
Watson and Walker

IBM

Enterprise2014

Why is the CPU Time For a Job so Variable?

Session zPE008

9.0

Welcome

- Hi, thanks for coming.

- Who I am and what we do.
  - Quarterly subscription newsletter – *Cheryl Watson's Tuning Letter*
  - Classes on z/OS new features, WLM, performance, software pricing, sysplex, and high availability
  - Consulting on all the above topics
  - Software products – GoalTender and Boxscore

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 presentations with the same title available on our website under 'Presentations' at www.watsonwalker.com.

Agenda

- Finding a sound basis for comparison
- Why This Topic?
- Recent hardware changes
- Recent customer experiences
- Hardware Changes Affecting Variability
- Software Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?
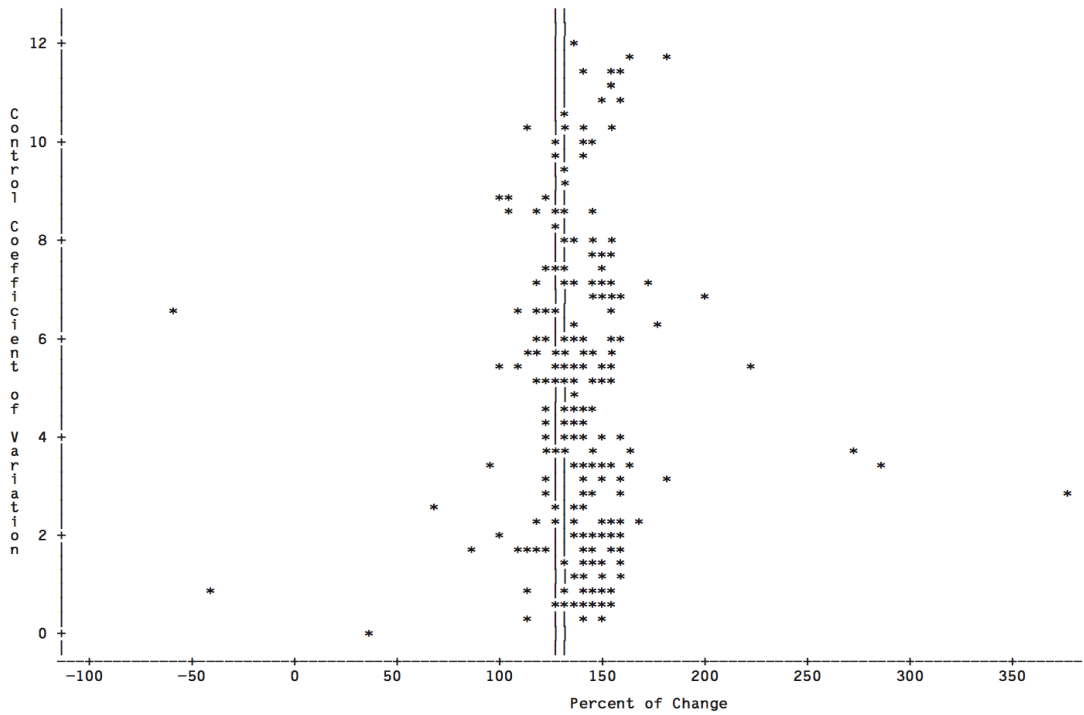
- PLEASE ask questions as I go along.

Finding a sound basis for comparison

- Ways to benchmark jobs:
  - IBM works VERY hard to develop an environment and processes to enable them to get repeatable measurements that can then be used to benchmark new CPCs.
    - For more information, see https://www-304.ibm.com/servers/resourcelink/lib03060.nsf/pages/lsprindex/$file/SC28118717.pdf
  - Ideally, take measurements at 90% or 99% busy, with all other work being the same (like IBM uses to come up with ITRRs – see pg. 21 of LSPR guide)
  - Select jobs that run after every change using identical data; problems occur because of other loads on the system
  - Even better would be to find every stable job step or transaction and look at the change in the CPU per I/O (take total CPU time and divide by the number of EXCPs) between the two environments; this gives you a "benchmark" that is truly representative of all your workload.

- First plot
  - BoxScore report showing a CPU upgrade that zPCR indicated should deliver a CPU speed increase of 132.6%. However, based on analysis of all the work that ran before and after the upgrade, they only observed 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?
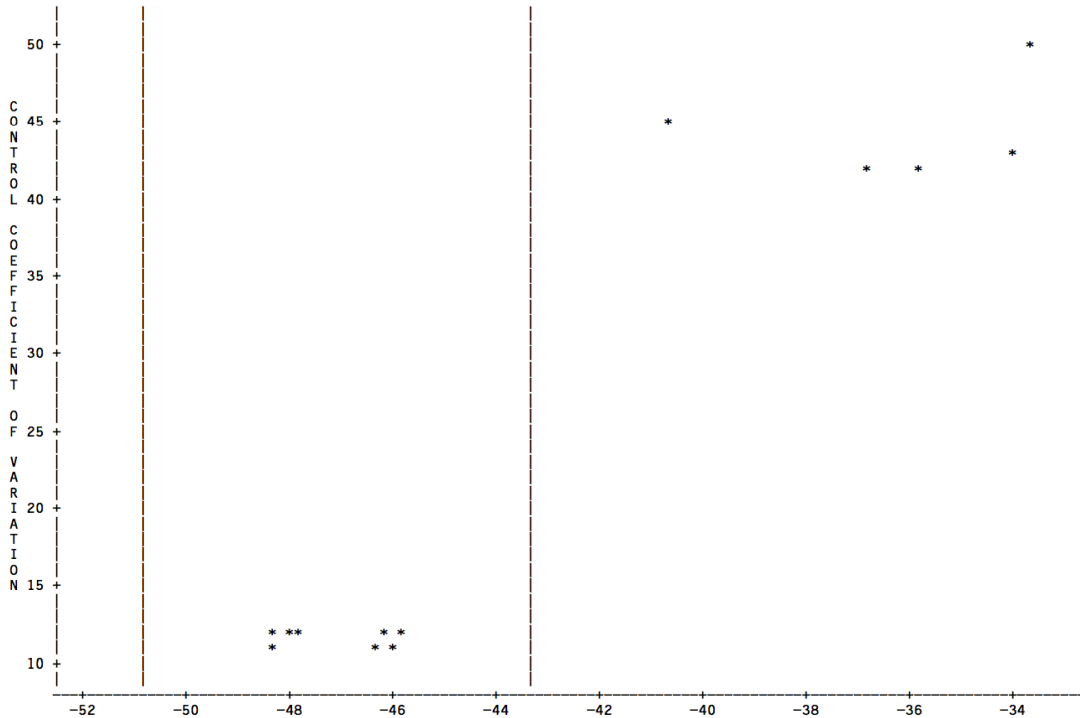   - None – some customers would be happy; some not so!

CPU Per I/O

BOXSCORE/IMS   9:15 MONDAY, September 2, 2013

PERCENT OF CHANGE — OBSERVED = −43.4, EXPECTED = −50.8

PLOT OF CV*PERCENT. SYMBOL USED IS '*'.

Agenda

- Finding a sound basis for comparison
- Why are people interested in this topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- Software Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

Why are people interested in this topic

- Recent Hardware Changes

  - IBM usually announces a new processor generation/family every year, alternating between the business class and enterprise class models

  - The average customer upgrades their processor every two to four years

  - Due to the amount of effort and cost, most customers wait for a CPC generation upgrade to also upgrade channels, network, Coupling Facilities, memory, and to add new features. The result is that *many* things (not just the CPUs) are being changed at one time

- Recent Hardware Changes
  - The recommended IBM processor sizing tool - 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 except for the CPU itself.
    - This is as it should be – otherwise, people that don't upgrade all those things would find that the new processor is slower than anticipated. Would you prefer to get MORE than you expected or LESS??!!

Why are people interested in this topic

- Recent Customer Experiences
  - Many upgrades from z9 to z114 or from z10 to z196, have provided even better savings (i.e. more capacity) than indicated by zPCR
  - Many moves from z9 or z114 to zBC12 or from z10 or z196 to zEC12, have provided more capacity than indicated by zPCR
  - The differences have been really dramatic – up to 35% (even 40% in one case) divergence from zPCR numbers
  - Outsourcers are being hurt by a drop in revenue. IT departments funded by chargeback are under-recovering. Customers (both outsourced and internal) are losing faith in charging/sizing mechanisms. *Nobody* understands what's happening.

- Recent Customer Experiences
  - People need to understand this stuff:
    - Because accurate capacity planning/upgrade decisions depend on understanding how IBM arrive at their numbers and how those numbers relate to YOUR environment
    - Because chargeback/cost recovery mechanisms should be equitable to both parties – provider and user
    - To arm you with answers when people complain about their bills
    - Because if you have outliers, you should understand why
    - Because everyone is under so much pressure to control costs – you can't run your configuration at maximum efficiency if you don't understand the capacity usage of your applications.
    - How do your chargeback algorithms handle uncaptured time or system "overhead"? Hardware or software or configuration changes can impact these so you need to be equipped to handle this.

Agenda

- Creating a sound basis for comparison
- Why are people interested in this topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- Software Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

- Typical changes that occur with each new generation
  - 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, on early CMOS machines, some COBOL programs took many times longer than expected; problem was tracked down to programs using subscripts instead of indexes. The CVB and CVD instructions had been moved from the hardware into microcode.

- zIIPs/zAAPs (specialty processors or SPs)
  - Run at full speed even if running on knee-capped, or sub-capacity CPCs (i.e. CP on zEC12 4xx is about 16% the speed of the 7xx, but a zIIP or zAAP on the 4xx is the same speed as the 7xx CP)
  - A job might run one day using CPs and SPs and the next day using only CPs; CPU time will differ; are your reporting tools set up to "correctly" handle all the SMF fields associated with this situation?
  - Slight CPU overhead in switching to an SP, but could be reduced CPU if on sub-capacity CPC, cost savings often seen in software and hardware pricing

- zEC12 (as an example of how new processor capabilities can impact your reports and bills)
  - 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 re-compiled with the new COBOL compiler
  - Pageable 1MB and fixed 2 GB Page frames
    - May reduce CPU time for users of DB2 buffer pools and Java heap
    - Fixed 1MB frames (available on earlier CPCs) definitely save CPU time
  - Decimal floating point zoned conversion facility can reduce CPU time for jobs compiled with the new PL/I compiler
  - Lookahead instruction paths can reduce CPU time of a job

Hardware Changes Affecting Variability

- zBC12 and zEC12
  - New Integrated Firmware Processor (IFP) processors move processing off the CP
    - Used for native PCIe functions such as zEDC Express and 10GbE RoCE Express
    - Great. However, they are not free to buy, yet their use is not reported in Type 30 records. How will this impact your chargeback? They cost YOU money to purchase, but they REDUCE your income…
  - Newer faster FICON channels (FICON Express8S)
    - Improved channel speed may reduce the CPU time of a job
  - OSA 10 GbE, 1000BASE-T
    - Improved network speed may reduce the job CPU time
  - Compression
    - Old style compression saves DASD space but eats CPU. PKZIP compression provides option to use a zIIP for the compression processing. zEDC moves the processing to an IFP.

- zBC12 and zEC12
  - More memory
    - More memory can 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 might result in less CPU time for paging operations

Hardware Changes Affecting Variability

- In the past, Moore's Law meant that much of the speed increase from one generation to the next was the result of changing chip technology.

- However, given the much-publicized ending of Moore's Law, it is likely that performance improvements in the future will be increasingly be delivered by enhancements (like transactional execution) that only affect a subset of programs

- This means that the challenges that people are facing today will only get worse in the future.

  - See Bob Rogers IBM Systems Magazine articles about SMT - http://www.ibmsystemsmag.com/mainframe/trends/IBM-Research/smt_mainframe/

- HiperDispatch
  - If turned on, CPU time of many jobs can be reduced (from 0% to 10%)
  - The benefit of HiperDispatch changes with each new hardware generation
    - On z10 (first generation to support HiperDispatch), the default setting for HD was NO, and the recommendation was to enable it on larger models.
    - On z196 (and later), the recommendation (and the default) is that HiperDispatch should be set to YES on all z/OS LPARs.
  - The effect of poorly configured LPARs (with a too-high logical-to-physical CP ratio) is minimized with HD
  - Alain Maneville (IBM France) has a very useful spreadsheet-based tool to help you plan HiperDispatch configurations – we will document it in a future Tuning Letter

- 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, local vs remote) – SHOULD reflect an informed trade off between performance, cost, and availability (single point of failure)
    - But this is not always the case…………
  - CPU time accounting for asynchronous requests (typically resulting from long response times) is not as accurate as for synchronous requests
    - Nearly all CPU time for a sync request is charged to requester.
      - Capture ratio for a test workload doing ONLY sync requests was 98%
    - For async requests, a lot of the CPU time is charged to XCFAS, some is charged to various other address spaces, and a LOT is uncaptured.
      - Capture ratio for a test workload doing ONLY async requests was 63%

Hardware Changes Affecting Variability

Systems and Technology Group

IBM

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

Courtesy of Gary King of IBM

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

Agenda

- Creating a sound basis for comparison
- Why are people interested in this topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- Software Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

Software Changes Affecting Variability

- Operating system levels affect CPU time (each new z/OS release generally uses less CPU to do the same amount of work).
  - Download IBM zSoftCap tool to get a projection of relative capacity requirements of different releases - https://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS268
- New z/OS features (e.g. PLAREA (automatic) and LFAREA(optional)) can impact subsystems
- New subsystem versions can automatically move work from CPs to zIIPs (DB2 V11, for example)
- Applying maintenance may introduce performance APARs that will affect the amount of CPU time consumed, especially if these are new function APARs.
  - Can also introduce bugs that consume MORE CPU
- Maintenance of vendor products may introduce changes in CPU times.

Agenda

- Creating a sound basis for comparison
- Why are people interested in this topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- Software Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

Environment Changes Affecting Variability

- CPU time of jobs and transactions are affected by total CPC utilization. An increase of 10% in physical CPU busy can increase job CPU time by 3-5%.
  - See White Paper titled "Running IBM System z at High CPU Utilization" by Gary King - https://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101208

- 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 based on the numbers 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*)
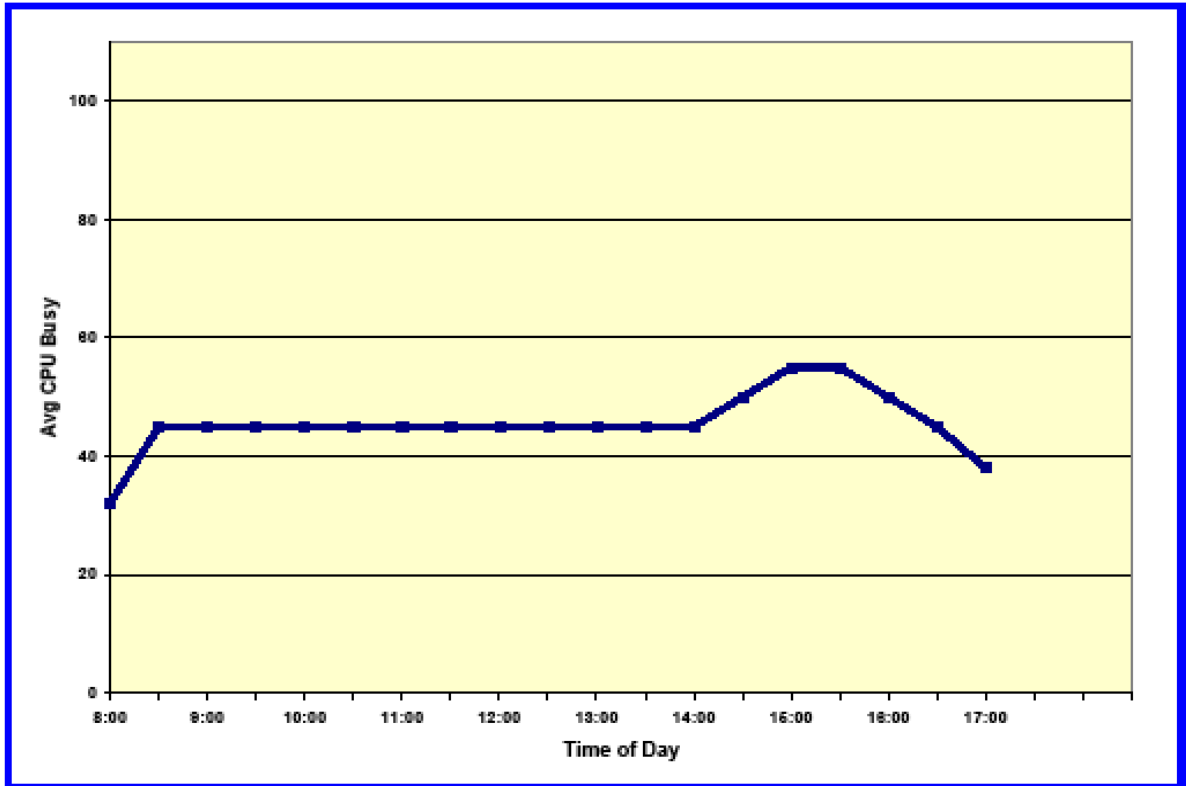
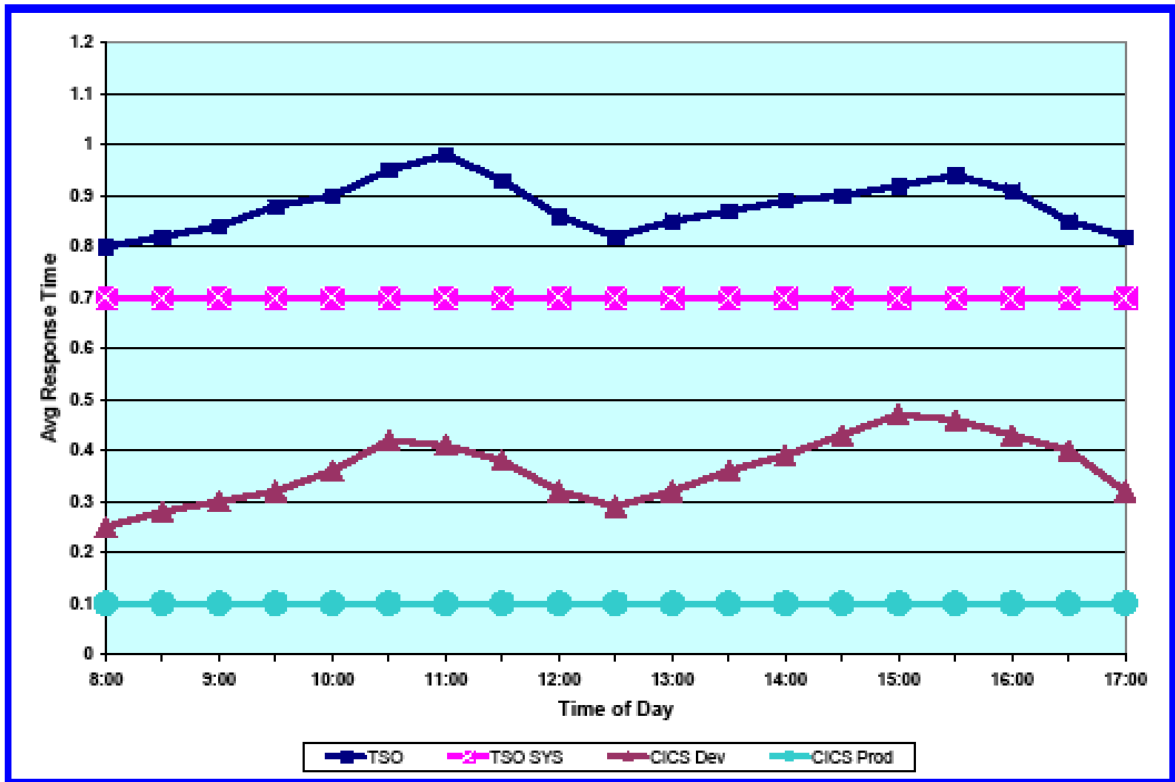# Environment Changes Affecting Variability
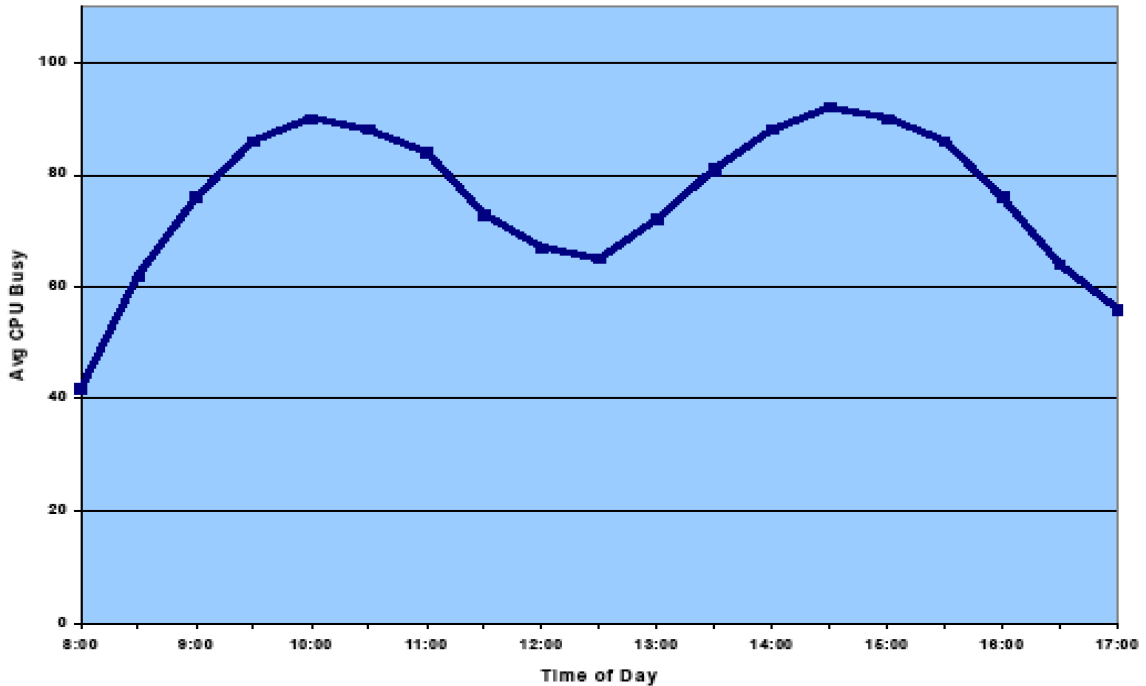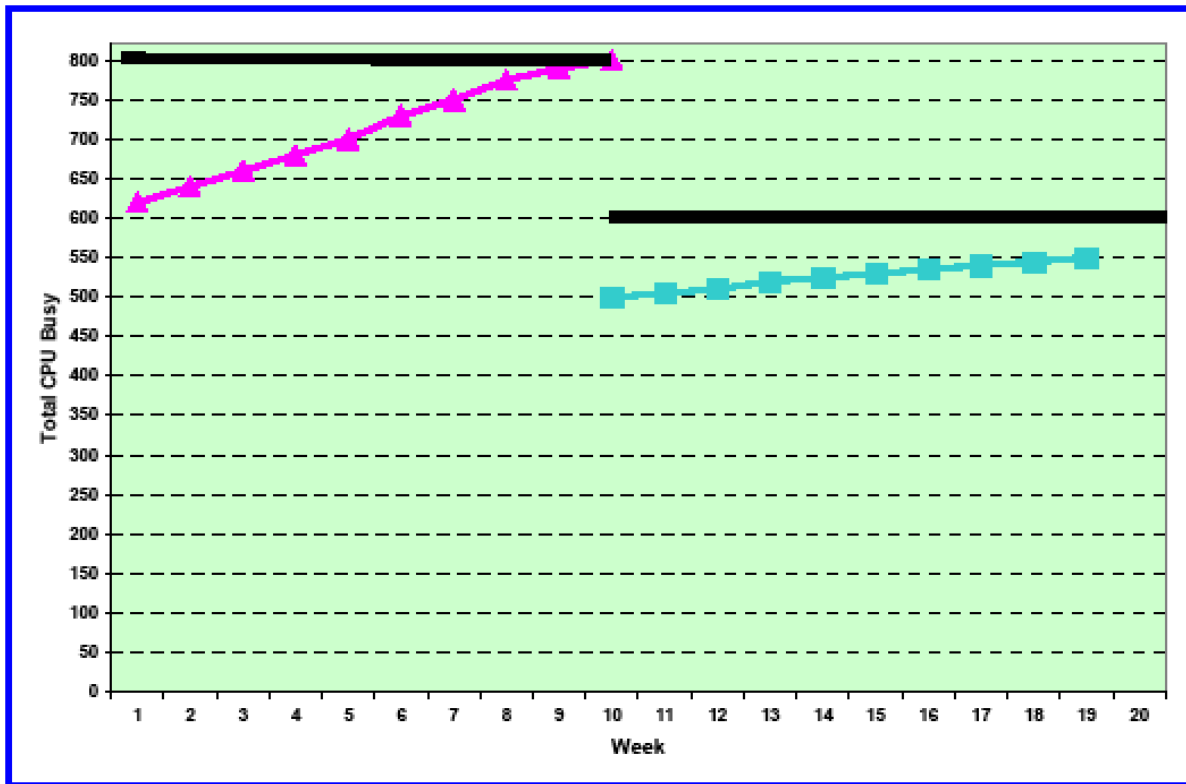
Environment Changes Affecting Variability
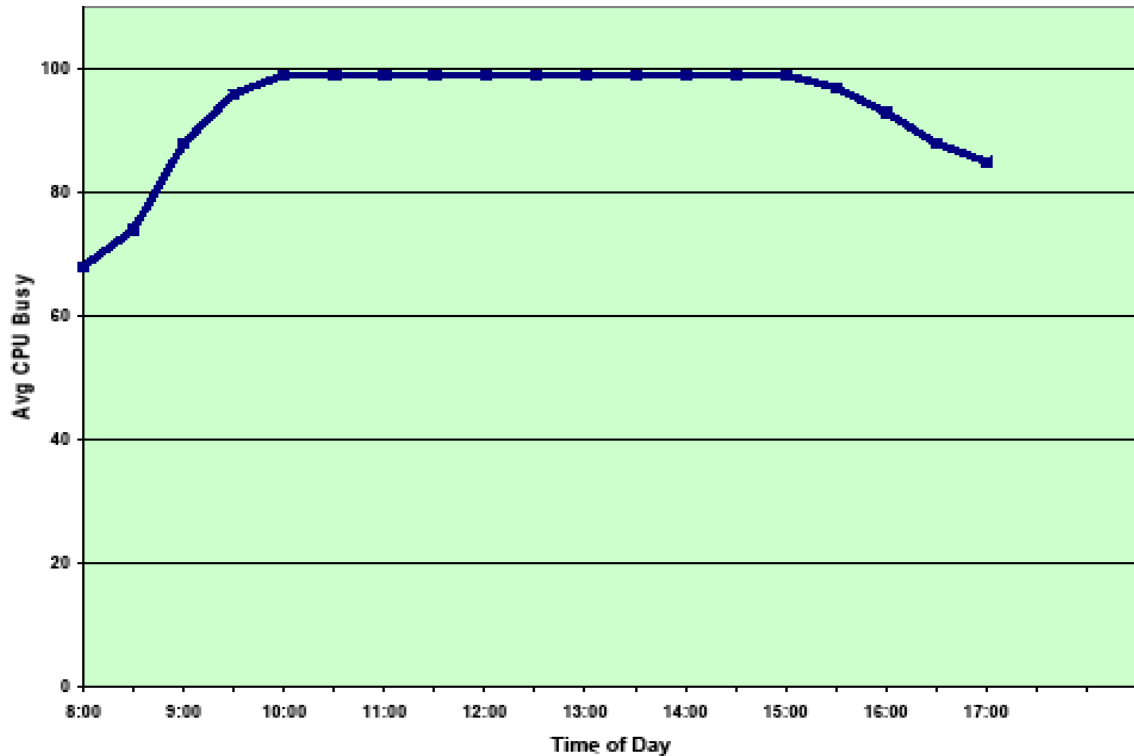
Environment Changes Affecting Variability

Environment Changes Affecting Variability

Environment Changes Affecting Variability

Environment Changes Affecting Variability

- Increase of LPAR weight can reduce job CPU time – more weight might result in LPAR getting more Vertical High CPs, so runs more efficiently. Conversely, decreasing weight can *in*crease job CPU time.

- Changes in workloads in other LPARs, change to the number of LPARs or the number of LPs in all LPARs can also change a job's CPU time.
  - Any time you make a change to something that uses shared resources, you need to consider the impact on other sharers.

Agenda

- Creating a sound basis for comparison
- Why are people interested in this topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- Software Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

Other Changes Affecting Variability

- Application changes can change CPU times.
  - Have you considered using changes in application CPU/IO to audit change management processes?
- 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
- Recompiling a program to apply a small change can dramatically impact CPU times depending on compiler levels and options
  - But consider DR site capabilities
  - MIGHT be worth re-compiling real CPU hogs with current compilers
- Number of interrupts on the system
- Room temperature
- "I didn't change anything!" changes

Agenda

- Creating a sound basis for comparison
- Why are people interested in this topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- Software Changes Affecting Variability
- Environment Changes Affecting Variability
- Other Changes Affecting Variability
- Which Measurements?
- What To Do?

- How do you determine what to expect?
  - Sites SHOULD be using zPCR to determine expectations.
    - Make sure you use Type 113 records to increase accuracy of zPCR predictions
  - 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
```

- 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/lsprindex?OpenDocument
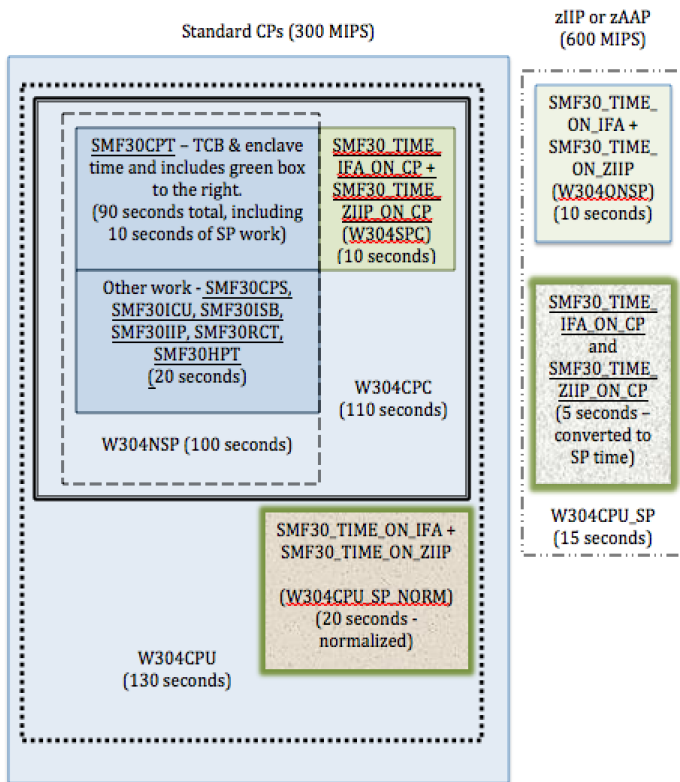 - *Cheryl Watson's System z CPU Chart*

- 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.

Measurements

•

Standard CPs (300 MIPS)

zIIP or zAAP (600 MIPS)

SMF30CPT – TCB & enclave time and includes green box to the right. (90 seconds total, including 10 seconds of SP work)

SMF30_TIME_IFA_ON_CP + SMF30_TIME_ZIIP_ON_CP (W304SPC) (10 seconds)

Other work - SMF30CPS, SMF30ICU, SMF30ISB, SMF30IIP, SMF30RCT, SMF30HPT (20 seconds)

W304CPC (110 seconds)

W304NSP (100 seconds)

SMF30_TIME_ON_IFA + SMF30_TIME_ON_ZIIP (W304QNSP) (10 seconds)

SMF30_TIME_IFA_ON_CP and SMF30_TIME_ZIIP_ON_CP (5 seconds – converted to SP time)

W304CPU_SP (15 seconds)

SMF30_TIME_ON_IFA + SMF30_TIME_ON_ZIIP (W304CPU_SP_NORM) (20 seconds - normalized)

W304CPU (130 seconds)

© Copyright Watson & Walker 2014

Agenda

- Creating a sound basis for comparison
- Why are people interested in this topic?
  - Recent hardware changes
  - Recent customer experiences
- Hardware Changes Affecting Variability
- Software 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, both expected and unexpected ("I didn't change anything – why are those jobs suddenly using twice as much CPU?").
- 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 tiering.

What to do?

- SMF Type 30, Counter Section
  - Activated when SMF30COUNT 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 investigating repeatability of these numbers. If you would like to contribute some SMF data, contact John Burg (jpburg@us.ibm.com).

Other sessions

If you found this session helpful, there are a couple of other sessions that you might be interested in:

– zPE007 – *The Skinny on Coupling Thin Interrupts* – **Frank Kyne** – Wed 13:00

– zHA001 - *Long Distance IBM Sysplex Data Sharing* - **Frank Kyne** – Thurs 13:00

Also, if you like SMF data (and who doesn't??!!), please see our new AND IMPROVED(!) *SMF Reference Summary* at www.watsonwalker.com/references.html
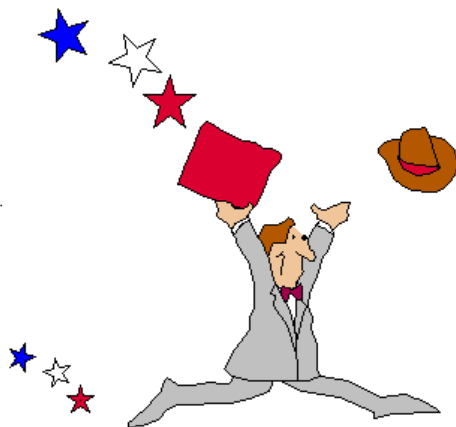
If you are interested in the financial aspects of running z/OS, Cheryl and I are running a class on Software Pricing Strategies in Dallas in December – we would love to see you there. And please bring a friend.

Any questions?
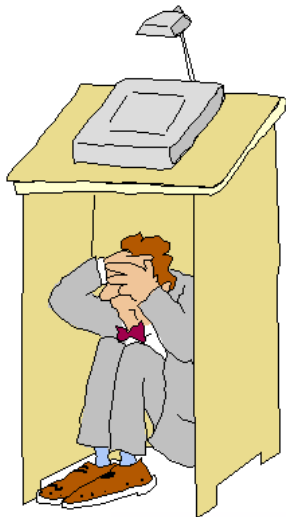
Thank you for coming and have a safe trip home.

Please remember to complete an evaluation Session number is zPE008

> **Meet the authorized IBM Global Training Providers in the Enterprise Solution Showcase**

- Access to training in more cities local to you, where and when you need it, and in the format you want
  - Use IBM Training Search to locate training classes near to you
- Demanding a high standard of quality / see the paths to success
  - Learn about the New IBM Training Model and see how IBM is driving quality
  - Check Training Paths and Certifications to find the course that is right for you
- Academic Initiative works with colleges and universities to introduce real-world technology into the classroom, giving students the hands-on experience valued by employers in today's marketplace
- www.ibm.com/training is the main IBM training page for accessing our comprehensive portfolio of skills and career accelerators that are designed to meet all your training needs.

Global Skills Initiative

ARROW

Avnet Academy
Global Training Provider

AVNET

Global Knowledge.

INGRAM MICRO

LearnQuest

Related learning: Performance track

- Courses:
  - AN51G: Power Systems for AIX IV: Performance Management
  - AN52G: Advanced Tools for AIX Performance Analysis
  - OL23G: IBM i Structure, Tailoring and Basic Tuning
  - OL66G: Performance Analysis and Capacity Planning for IBM i
  - ES42G: Parallel Sysplex Implementation Workshop
  - ES73G: IBM System z Parallel Sysplex Operations
  - ES85G: Advanced z/OS Performance: WLM, Sysplex, UNIX Services and Web
  - ES90G: Advanced Parallel Sysplex Operations and Recovery Workshop

- Training paths:
  - IBM Power Systems and AIX
  - AIX Users and System Administration
  - ✓ **AIX Security, Network Administration, and Performance** *(track relevant)*
  - Power Systems for AIX Virtualization
  - PowerHA SystemMirror for AIX
  - Cloud on Power
  - IBM i: End User
  - IBM i: System Operator
  - ✓ **IBM i: System Administrator** *(track relevant)*
  - IBM i: Application Developer
  - IBM i: Virtualization
  - IBM i: Solutions

  - IBM i: WebSphere Application Server Administrator
  - IBM i: Networking
  - AIX Systems Management, Clustering, Internals, and Cloud Computing
  - z/OS and System z
  - ✓ **z/OS Operation and Parallel Sysplex** *(track relevant)*
  - z/OS Installation, Tuning, JES, and REXX Programming and Problem Diagnosis
  - Security Skills
  - z/OS UNIX System Services, WebSphere and SAP
  - z/VM and Linux on System z
  - z/OS Networking