

Cheryl's Hot Flashes

FLCMG
Tampa, Florida
January 18, 2011

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

Cheryl Watson's Tuning Letter, CPU Charts, **BoxScore** and **GoalTender**

Agenda

- ✦ zEnterprise Observations
- ✦ *If I Ran* the Site
- ✦ User Experiences
- ✦ Interesting Notes

zEnterprise Observations



- ✦ In my opinion, the July 22nd announcement was the most significant in the last decade. (As an aside – I've been watching IBM announcements since 1965.)
- ✦ Realize that there are really three parts of a zEnterprise –
 - ✦ z196 – faster processor, better memory management, 100 new instructions, new pricing, lower cost specialty engines
 - ✦ This is the piece that most installations will see in the next couple of years.
 - ✦ zBX – blade servers closely linked to z196/z10s. Initial benefit comes from reduction in network costs/response, smaller footprint, reduction in energy, reduction in TCA, reduction in people costs.
 - ✦ URM – Unified Resource Manager. zManager has potential.

zEnterprise Observations



- ✦ Biggest benefit will come from analytic processing and new industry solutions.
 - ✦ How about a 4-hour DB2 query completing in just seconds?
 - ✦ See the IBM Smart Analytics Optimizer for DB2 on z/OS SHARE sessions.
- ✦ Okay, I'm sold, but what aren't you seeing? What questions should you be asking? What does it mean for the majority of sites today?
 - ✦ All of IBM's benefits (of course, but you've heard these)
 - ✦ Lower specialty processor costs
 - ✦ Chargeback and capacity planning complexity with zBXs
 - ✦ MIPS have changed

z196 Specialty Engines - IFL

Status	U.S. Price	Avg RNI MIPS (single IFL)	\$ per MIPS
New z9-BC	\$95,000	459	\$207
New z9-EC	\$125,000	560	\$223
New z10-BC	\$75,000	661	\$113
New z10-EC	\$75,000	901	\$83
New z196	\$55,000	1204	\$46
z9-EC to z10-EC Upgrade	\$0	+341	\$138
z9-EC to z196 Upgrade	\$33,000	+644	\$131
z10-EC to z196 Upgrade	\$16,500	+303	\$75

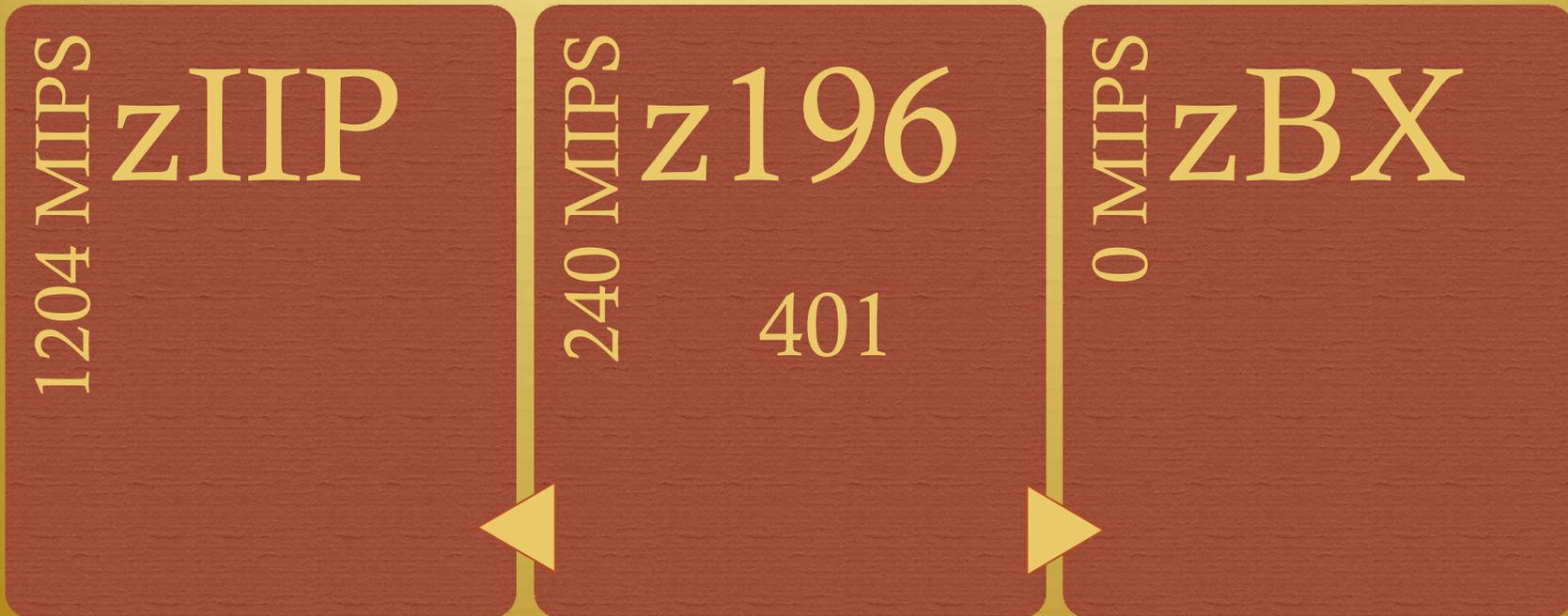
MIPS estimates come from *Cheryl Watson's CPU Charts, July 29, 2010*

z196 Specialty Engines – zIIP/zAAP

Status	U.S. Price	Avg RNI MIPS (single IFL)	\$ per MIPS
New z9-BC	\$95,000	459	\$207
New z9-EC	\$125,000	560	\$223
New z10-BC	\$125,000	661	\$189
New z10-EC	\$125,000	901	\$139
New z196	\$100,000	1204	\$83
z9-EC to z10-EC Upgrade	\$0	+341	\$139
z9-EC to z196 Upgrade	\$60,000	+644	\$153
z10-EC to z196 Upgrade	\$30,000	+303	\$128

MIPS estimates come from *Cheryl Watson's CPU Charts, July 29, 2010*

z/OS zBX Issues



z/OS zBX Issues

- ✦ When you do capacity planning and chargeback for DB2, Java, etc., how do you deal with CPU time accumulated on a zIIP or zAAP?
- ✦ Example: A CICS transaction issues a request to DB2, which runs part of the query on a zIIP.
 - ✦ Some days it runs on the zIIP, some days it doesn't.
 - ✦ When it all runs only on the z10, it accumulates CPU time at speed of z10.
 - ✦ When part of it runs on the zIIP, it accumulates CPU at speed of zIIP (can be several times faster than z10).
 - ✦ The SMF records contain a normalization factor so you can charge the zIIP time at native speed or normalized speed. It's your option.

z/OS zBX Issues



- ✦ Now add a zBX
 - ✦ Using same example, assume that DB2 and IBM Smart Analytics Optimizer decide to send the query to a blade server.
 - ✦ The zBX does not have MIPS assigned, so there is no normalization factor.
 - ✦ One day it could cost a user \$10 for a request if it used the zBX, but the next day it might cost \$500.
 - ✦ We don't yet know what metrics will be available for the time spent on the zBX.
 - ✦ How will you charge? How will you plan the capacity?

z/OS 1.11 LSPRs

- ✦ Large System Performance Reference (LSPR)
- ✦ Location of LSPR website has changed - <https://www-304.ibm.com/servers/resourceLink/lib03060.nsf/pages/lspindex?OpenDocument>
- ✦ July 22, 2010 – New LSPRs published
 - ✦ Software stack is now z/OS 1.11 versus z/OS 1.9.
 - ✦ Workloads are based on Relative Nest Intensity (RNI) as low, average, and high versus older workloads of CB-L, ODE-B, WASDB, OLTP-T, OLTP-W, MIX, LoIO-Mix, and TI-Mix.
 - ✦ RNI can be estimated using values in the SMF Type 113 records.
 - ✦ Most installations found that LoIO-MIX came closest to their own workloads.
 - ✦ IBM Publishes PCI (Processor Capacity Index) – same values as average MIPS.
 - ✦ PCI has been “adjusted” to more closely match the LoIO-MIX.
 - ✦ zPCR will be updated in mid-August with new LSPRs.

z/OS 1.11 LSPRs



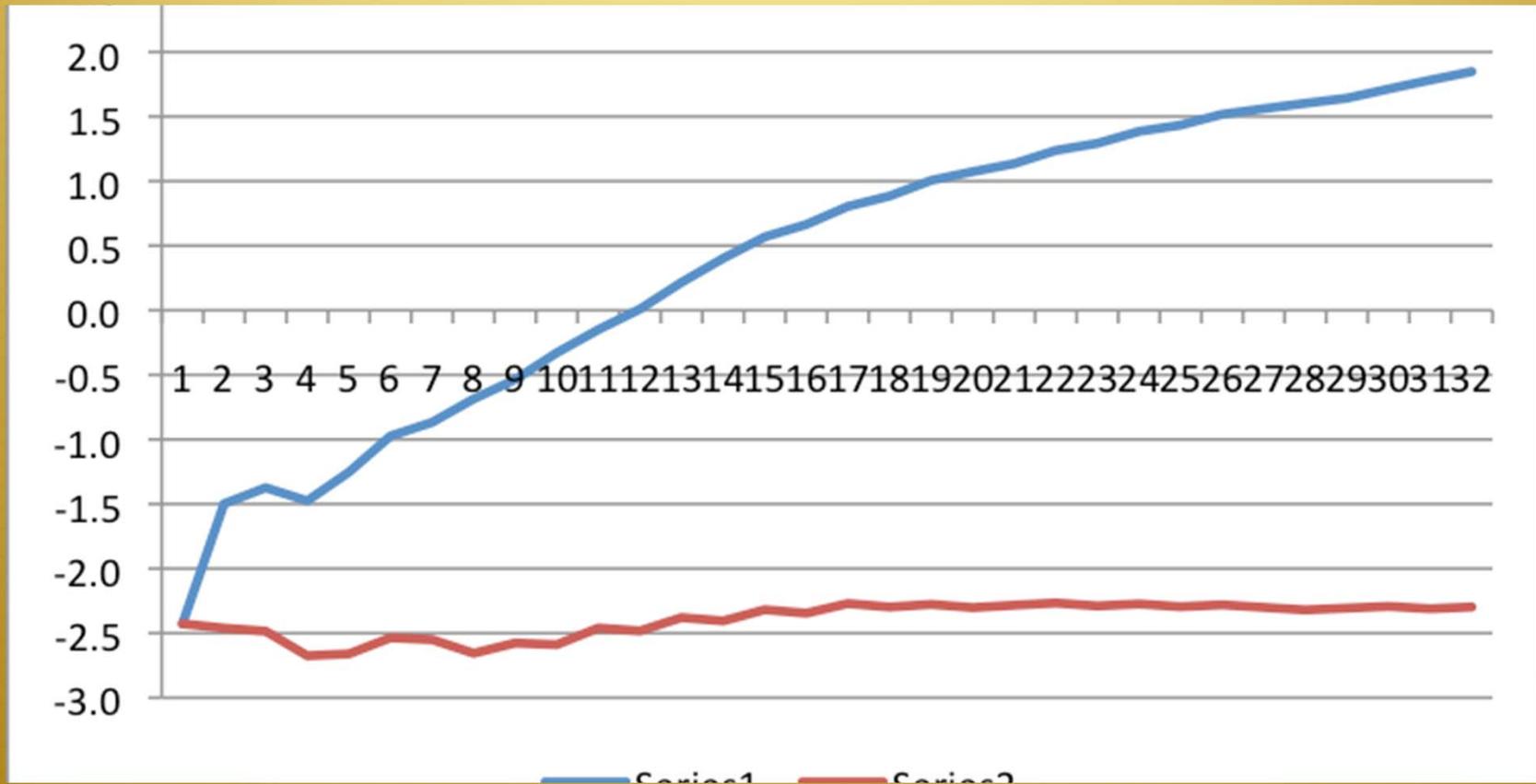
- ✦ Most important points!
 - ✦ Run CPU MF and HIS in order to collect SMF type 113s
 - ✦ Process 113s to determine RNI workload
 - ✦ Run zPCR to determine expected capacity
 - ✦ Should also use this for any type of change – adding, deleting CPs, LPs, LPARs, etc.

z/OS 1.11 LSPRs



- ✦ Caution – your MIPS may change
- ✦ The next slide shows the percent change in MIPS between the MIPS derived from z/OS 1.9 LSPRs and the z/OS 1.11 LSPRs for the first 32 processors.
- ✦ Notice that “average MIPS” are lower on the smaller z/OS 1.9 LSPRs and higher on the larger z/OS 1.9 LSPRs.
- ✦ Notice that the “average RNI MIPS” based on z/OS 1.11 are about 2.5% lower than the Low I/O MIPS, but that’s pretty even across all sizes.
- ✦ If you used Low I/O in the past, then new MIPS will average about 2.5% lower for the z10-EC with 1 to 32 CPs

z/OS 1.11 LSPRs



zEnterprise Pricing



✦ z196 Software Prices

- ✦ There is no “tech discount” obtained from a reduction in MSU assignments for the z196.
- ✦ In the past, there was up to a 10% reduction in MIPS/MSU, but that is no longer true for the z196.
- ✦ The MIPS/MSU remains the same.
- ✦ But new pricing method, AWLC, is now available for zEnterprise machines.
- ✦ VWLC pricing above 575 MSUs has no discount for subsystems such as DB2, CICS, IMS, etc. sites).
- ✦ AWLC pricing provides discounts up to about 2,000 MSUs.

If I Ran the Site

“It’s a pretty good site,”
Said young Sally Ann White,
“and the people who run it
Are really quite bright.”



“But if I ran the site,
Said young Sally Ann White,
“I’d make a few changes.
And I’d do it right”

With thanks to Dr. Seuss.

If / *Ran* the site . . .

- ✦ While attending SHARE in Boston, I kept seeing suggestion after suggestion on ways to improve the data center. Many of these are the same recommendations that have been given for years.
- ✦ Why aren't people doing them?
 - ✦ Not enough staff.
 - ✦ It takes so much time to install a release, roll it out, and maintain it, that there's no time left to exploit it.
 - ✦ Managers don't want any changes that could impact production.
 - ✦ Too many inexperienced people.
 - ✦ Internal politics.
 - ✦ Green screen is faster.

If / *Ran* the site . . .



- ✦ I'd start with baby steps, such as:
 - ✦ Schedule an hour a week, an hour a day, one day a month to work on enhancements.
- ✦ The next few slides show some of the areas I would start with (most assume that you're on z/OS 1.11).
- ✦ If not on z/OS 1.11, then start planning for z/OS 1.12. I would rather be on 1.12 than on 1.11 because of added performance benefits and features.
- ✦ Do all of these on your test system first!
- ✦ There is no specific priority or order here.

If / *Ran* the site . . .

- ✦ Automate as much as possible (that is, let the system do some of your work!)
 - ✦ Turn on and use IBM Health Checker (z/OS 1.10+)
 - ✦ See SHARE sessions from Peter Relson, Paul Robichaux
 - ✦ See Seattle SHARE session 2240
 - ✦ <http://www.ibm.com/systems/z/os/zos/hchecker/>
 - ✦ Start using AUTOIPL (z/OS 1.10+)
 - ✦ See SHARE sessions from Mark Brooks
 - ✦ Implement sysplex-aware DAE (Dump Analysis and Elimination)
 - ✦ SHARE papers are at www.share.org.

If / *Ran* the site . . .

- ✦ Automate (continued)
 - ✦ Exploit SMF Logger
 - ✦ See Denver and Seattle sessions
 - ✦ Consider CA's SMF Director product
 - ✦ Upload SMF data into different archive logs with different retentions
 - ✦ Some data needs to be kept several years (e.g. type 30s when doing chargeback).
 - ✦ Some data needs to be kept for only a few days or weeks (e.g. type 99s, type 74s, network records).
 - ✦ Some data is so large that you want only one program reading it, not everyone who has to look at SMF (e.g. DB2 101).

If / *Ran* the site . . .

✦ Automate (continued)

- ✦ Use both Operlog and Syslog
 - ✦ They serve different purposes and provide backup
- ✦ Turn on HiperDispatch
 - ✦ Why waste up to 10% of your MIPS?
 - ✦ See SHARE sessions from John Burg, Gary King, and Kathy Walsh
- ✦ Set MSO Service Definition Coefficient (SDC) to 0.0
 - ✦ I'll cover the reason later
- ✦ If you have z/OS CA products, run (don't walk) to get CA's Mainframe Software Manager (MSM)
 - ✦ It's free, and makes installation a breeze!
 - ✦ I've heard great feedback from customers

If / *Ran* the Site

- ✦ Automate (continued)
 - ✦ Implement Sysplex Failure Management (SFM) with BCPii
 - ✦ See SHARE sessions from Mark Brooks & Steve Warren
 - ✦ See Seattle SHARE sessions 2251 and 2227 (Mark Brooks, Riaz Ahmed, Steve Warren)
 - ✦ Implement a display of IPLINFO at startup or periodically
 - ✦ See Mean Time to Recovery (MTTR) Redbook – [SG24-7816-00](#)
 - ✦ Use IBM defaults wherever possible in Parmlib members (such as IEASYSxx, IEAOPTxx, etc.), because IBM changes the defaults for best practices (i.e. don't code the values)
 - ✦ Use system symbolics whenever possible to reduce versions of parmlib members

If / *Ran* the site . . .

✦ Exploit Newer Features

✦ Implement z/OSMF

- ✦ See SHARE sessions by Anuja Deedwaniya and Greg Daynes
- ✦ Yes, WAS OEM can be a hard install, so let someone new do it.
- ✦ If you tried before March, try again with new doc and scripts.
- ✦ z/OSMF 1.12 will provide a lot more benefit.
- ✦ I really believe this is our/your future!
- ✦ Start with simplifying problem determination.

✦ Migrate HFS systems to zFS systems

- ✦ It's an easy migration, and HFS is going away.

✦ Install and use Sysplex Timer Protocol (STP)

- ✦ z196 requires it. (You've had six years!)

If / *Ran* the site . . .

- ✦ Exploit Newer Features (continued)
 - ✦ Exploit IFLs, zIIPs, and zAAPs
 - ✦ Greatest benefit is on z196
 - ✦ Investigate all new pricing options, especially VWLC, eWLC, and now AWLC on z196
 - ✦ Proper alignment of LPARs could save millions!
 - ✦ Run the CPU Measurement Facility (MF) and the Hardware Instrumentation Services (HIS) address space continually
 - ✦ The measurements are done in the microcode, and the z/OS piece is really minimal cost
 - ✦ Collect the SMF type 113 records – send them to John Burg (especially before going to a z196)
 - ✦ See sessions from John Burg, Gary King, and Peter Enrico

If *I Ran* the Site



- ✦ Exploit New Features (continued)
 - ✦ Learn to use zPCR!
 - ✦ I mention this every SHARE, but more people need to use this (especially for z196)!
 - ✦ See previous “Hot Flashes” and search SHARE sessions from John Burg and Kathy Walsh
 - ✦ Exploit the Coupling Facility
 - ✦ Some sites still do minimal work on CF
 - ✦ Exploit it when possible
 - ✦ See presentations on Coupling Facility

If / *Ran* the Site



- ✦ Exploit New Features (continued)
 - ✦ Stay on most current CFCC level for your Coupling Facility
 - ✦ Don't get behind, or you may miss performance improvements
 - ✦ Use newest hardware when it's beneficial
 - ✦ Best example – be sure to replace ISC links on a z196 with Infiniband links
 - ✦ Several SHARE sessions on Infiniband
 - ✦ Next slide comes from two IBM presentations from Mark Brooks and Gary King on CFs. (Thanks so much.)

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

Host CF	z890	z990	z9 BC	z9 EC	z10 BC	z10 EC	z196
z890 ISC	13%	15%	16%	17%	19%	21%	NA
z890 ICB	9%	10%	10%	11%	12%	13%	NA
z990 ISC	13%	14%	14%	15%	17%	19%	NA
z990 ICB	9%	9%	9%	10%	12%	13%	NA
z9 BC ISC	12%	13%	14%	15%	17%	19%	23%
z9 BC PSIFB 12X	NA	NA	NA	NA	13%	14%	16%
z9 BC ICB	8%	9%	9%	10%	11%	12%	NA
z9 EC ISC	12%	13%	13%	14%	16%	18%	22%
z9 EC PSIFB 12X	NA	NA	NA	NA	13%	14%	16%
z9 EC ICB	8%	8%	8%	9%	10%	11%	NA
z10 BC ISC	12%	13%	13%	14%	16%	18%	22%
z10 BC PSIFB 12X	NA	NA	11%	12%	13%	14%	15%
z10 BC ICB	8%	8%	8%	9%	10%	11%	NA
z10 EC ISC	11%	12%	12%	13%	15%	17%	22%
z10 EC PSIFB 12X	NA	NA	10%	11%	12%	13%	15%
z10 EC ICB	7%	7%	7%	8%	9%	10%	NA
z196 ISC	NA	NA	11%	12%	14%	16%	21%
z196 PSIFB 12X	NA	NA	9%	10%	11%	12%	14%

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

PSIFB 1X links would fall approximately halfway between PSIFB 12X and ISC links

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

Mark & Gary's Notes

The coupling efficiency of a Parallel Sysplex cluster, particularly one that has heavy datasharing, is sensitive to the performance of the operations to the Coupling Facility. The chart estimates the “host effect” for a heavy data sharing production workload for various combinations of host processor and coupling technology. The values in the table represent the percentage of host capacity that is used to process operations to the coupling facility (which includes the time spent communicating with the CF and the time spent waiting for the CF to process the request). For example, a value of 10% would indicate that approximately 10% of the host capacity (or host MIPS) is consumed by the subsystem, operating system and hardware functions associated with coupling facility activity. The table is based on a “coupling intensity” of 9 CF operations per million instructions (MI), which is typical of high end data sharing work loads.

The values in the table can be adjusted to reflect the coupling intensity for any workload. One can calculate the coupling intensity by simply summing the total req/sec of the CFs and dividing by the used MIPS of the attached systems (MIPS rating times CPU busy). Then, the values in the table would be linearly scaled. For example, if the workload was processing 4.5 CF operations per million instructions (or 4.5 CF ops/second/MIPS), then all the values in the table would be cut in half.

For 9 CF requests/MI, host effect values in the table may be considered capped at approximately 18% due to z/OS Synchronous to Asynchronous CF Message Conversion. Configurations where entries are approaching 18% will see more messages converted to asynchronous. z/OS converts synchronous messages to asynchronous messages when the synchronous service time relative to the speed of the host processor exceeds a breakeven threshold at which it becomes cheaper to go asynchronous. When all CF operations are asynchronous, the overhead will be about 18%. By the time you have reached $\geq 18\%$ in the table, that corresponds to the time z/OS must have been converting almost every operation asynchronous. The 18% cap scales proportionally with the CF requests/MI activity rate. For example, at 4.5 CF requests/MI, the cap would 9%.

The hardware cost can be minimized by using the most efficient links with faster engine speeds for the CFs. This reduces the time that z/OS is waiting for the response while the message is on the CF link and while the CF is busy processing the request. With this in mind, it becomes obvious that the best coupling efficiency is generally seen when the CF is on the fastest processor and connected to the z/OS image via the fastest links. The chart bears this out. For example, holding the CF and host processor technology constant, the chart shows that coupling efficiency increases with faster links (z/OS spends less time waiting because the communication with the CF is faster). For a given host processor and link technology, coupling efficiency increases with faster CF technology (z/OS spends less time waiting because the CF processes the request faster). In most cases, upgrading to faster links has a more dramatic impact on coupling efficiency than upgrading to a faster CF.

If / *Ran* the Site

✦ Exploit New Features (continued)

- ✦ If you have RMF, use the Distributed Data Server (DSS)
 - ✦ This only needs to run on a single image in the sysplex
 - ✦ Collects RMF data from all images
 - ✦ Can be used to get sysplex view without running merged (and sorted) post-processor data
 - ✦ Is used by z/OSMF 1.12 for monitoring performance
 - ✦ Can be CPU-intensive if you aren't careful about cycle and interval times – turn off data that you don't need
- ✦ If you have z196, exploit FICON Discovery and AutoConfiguration (zDAC)

User Experiences

✦ DFSMSdss

- ✦ z/OS 1.8 to 1.10 (APAR OA25644) and z/OS 1.11 added CSI (Catalog Search Interface) filtering
- ✦ **OA32165** – While Doing Generic Filtering During COPY,DUMP, Using Filter of * as the First Char of First Qualifier Won't Find any DS's (OPEN)
- ✦ **OA32120** – New 1.11 CSI Filtering Causes Increase in Job Runtime (17Jun2010)
- ✦ **OA25644** – Catalog Filtering Enhancement (20Aug2009)
- ✦ Runtime increase only seen when using INCAT on the DSS command

- ✦ Thanks to Michael Friske, Fidelity Investments

User Experiences

- ✦ Problem: CICS response time was terrible on one region.
- ✦ Initial review saw that some timing fields were wrapping.
- ✦ WLM simulates a restart of the address space (this was happening several times a day)
- ✦ What caused the wrapping?
- ✦ Service unit accumulation:
 - ✦ OUCBCPU – 0754237D
 - ✦ OUCBIOC – 00028CD1
 - ✦ OUCBMSO – E87A0E93
 - ✦ OUCBSRB – 00358831
 - ✦ Due to CPU=1.0, IOC=0.5, MSO=0.0001, SRB=1.0
- ✦ Lesson – set MSO=0.0

User Experiences



- ✦ DB2 Lock Structure Duplexing Problem
 - ✦ Duplexing broke frequently
 - ✦ **OA31254** – IXC522I Break Duplex (27Jan2010)
 - ✦ This caused another problem:
 - ✦ Reestablishing duplexing on a DB2 lock structure caused a 5 to 10 second delay in the IMS apps.
 - ✦ Changing the structure definition from ENABLED to ALLOWED eliminated the problem when duplexing stops

Interesting Notes

- ✦ Peter Muench, IBM
 - ✦ SHARE Session – *RMF: The Latest and Greatest*
 - ✦ Showed RMF Enhancements, including
 - ✦ Enhanced Group Capacity Reporting
 - ✦ Monitor II OPT Report
 - ✦ Postprocessor XML Formatted Reports
 - ✦ HTTP API to Access Historical Data
 - ✦ Distributed Data Server (DSS) Report Selection
 - ✦ In-Ready Work Unit Queue Distribution
 - ✦ z/OSMF Resource Monitoring (**my favorite!**) – see next two slides from his presentation. Thanks, Peter

IBM z/OSMF Management Facility

Resource Monitoring – Sysplex Status Task

IBM z/OS Management Facility - Mozilla Firefox

https://boermf4.boeblingen.de.ibm.com:9443/zosmf/

IBM z/OS Management Facility Welcome bhbe Log out IBM

Welcome Monitoring D... Sysplex Status

Sysplex Status [Help](#)

Use this panel to quickly assess the performance of the workloads running on the sysplexes in your installation. You can also use this panel to define the sysplexes and Linux images that you want to monitor in the Monitoring Desktops task.

Resources

Actions ▾

Resource	Connectivity	Performance Index Status	Related Service Definition	Active WLM Policy
<input type="radio"/> LOCALPLEX	Connected	✓ PI <= 1 for all periods	RTDST3	RTDST
<input type="radio"/> SYSDPLEX	Connected	✗ PI > 1 for important periods	system	POLICY01
<input type="radio"/> SCLMPLEX	Connected	✓ PI <= 1 for all periods	Default	STANDARD
<input type="radio"/> TRXPLEX	Connected	✗ PI > 1 for important periods	Ralfpol	BASEPOL
<input type="radio"/> SHARPLEX	Connected	✓ PI <= 1 for all periods	SHARPLEX	SHAREPOL

Total: 5

Refresh Last refresh: 22.02.2010 17:57:24 local time (22.02.2010 16:57:24 GMT)

Automatic refresh

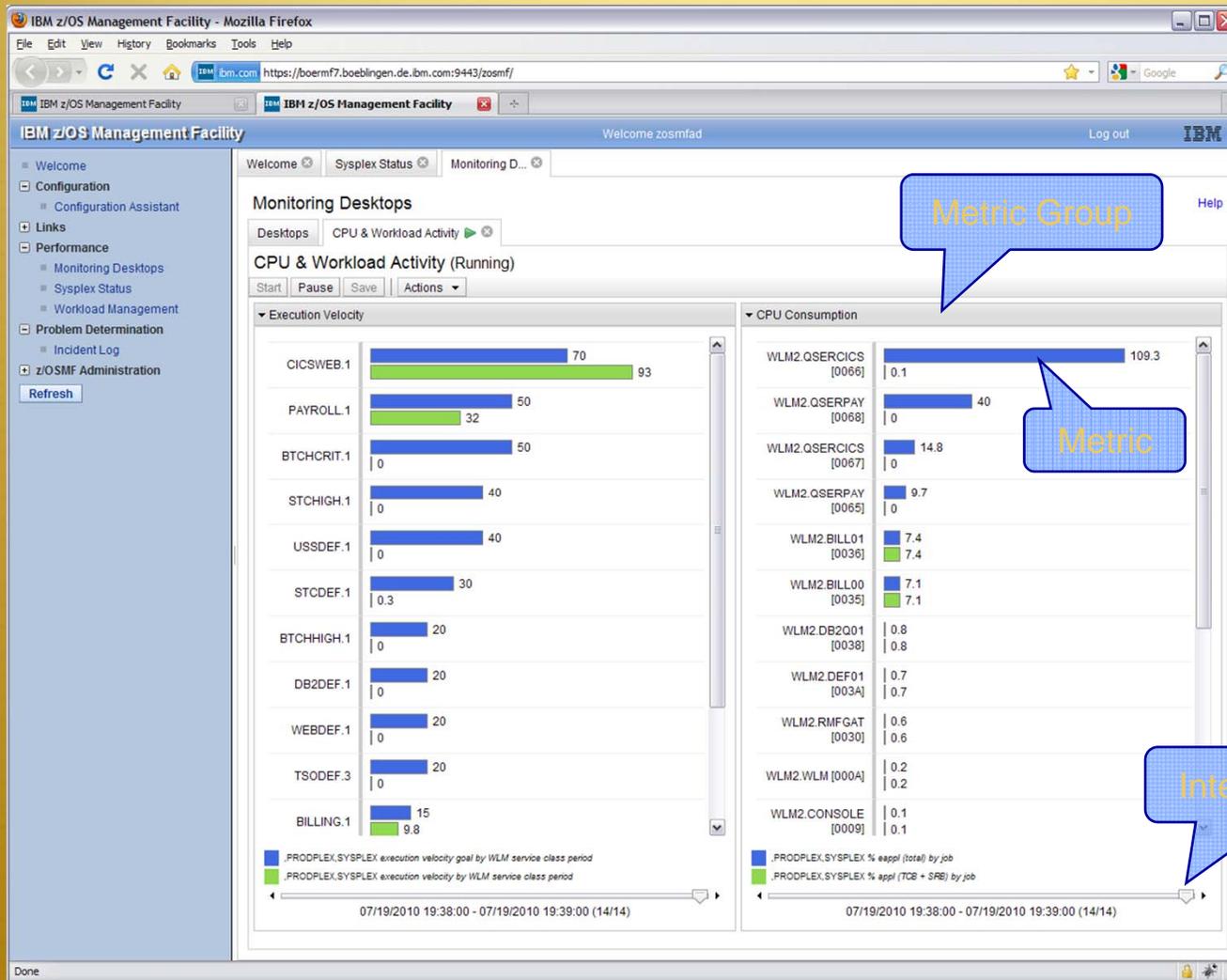
Fertig

Automatic refresh

Enterprise-wide Health check

Resource Monitoring — Monitoring Desktops

At a Glance



Interesting Notes



✦ Meral Temel, Garanti Technology

- ✦ See her SHARE session on *Performance Management Hints Using RMF data: Customer Experience*
- ✦ It's a big large (32 MB), but worth the download time
- ✦ See how performance can be automated

✦ Greg Daynes, IBM

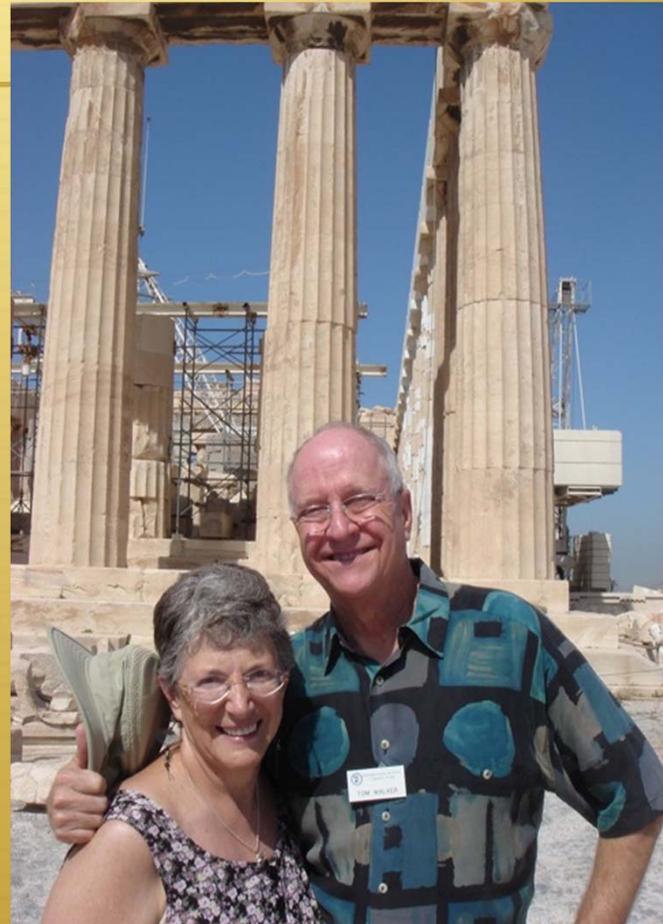
- ✦ See his SHARE session on *z/OS Software Positioning for the IBM zEnterprise System*
- ✦ 96 pages – this is a requirement if you're planning on a z196

Why Conferences?



- ✦ zEnterprise
- ✦ zBX
- ✦ z196
- ✦ z/OS 1.12
- ✦ Unified Resource Manager
- ✦ BladeCenter
- ✦ Blades
- ✦ Hypervisors
- ✦ Firmware
- ✦ Ensembles . . . My brain is full!

See You Later!



✦ Email: technical@watsonwalker.com

✦ Web site: www.watsonwalker.com

Cheryl Watson Walker with partner, husband, and best friend Tom Walker in Athens in June