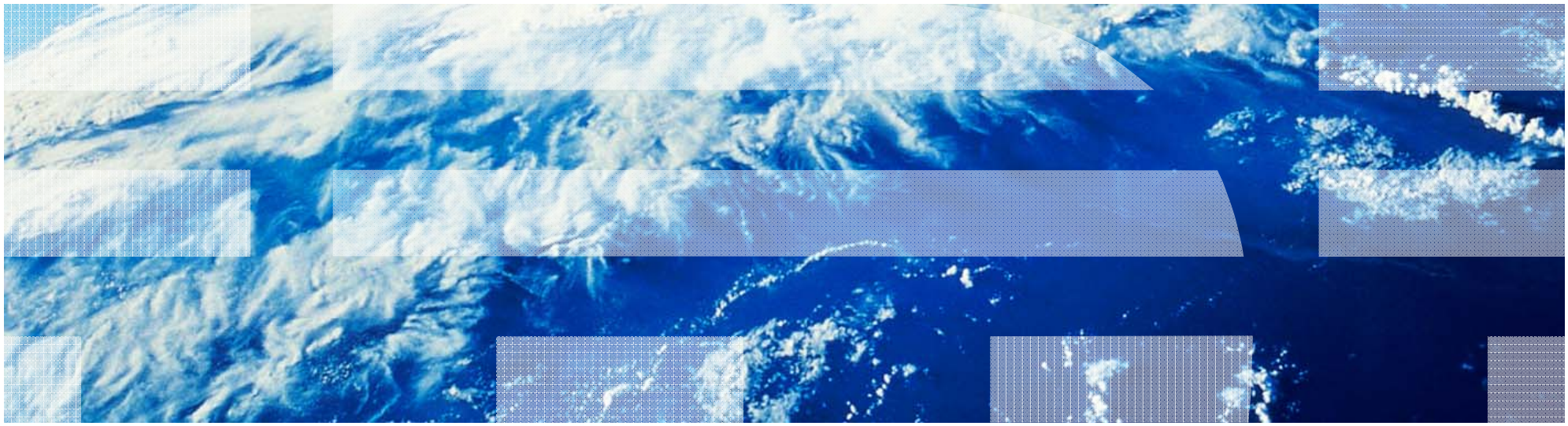
z/OS Performance **HOT** Topics

Session 9909

Kathy Walsh

IBM



Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:
Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.
IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.
All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.
This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.
All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.
Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Agenda

- Processor Information
  - IBM® zEnterprise™ 114 (zxxx)
  - CPU Measurement Facility
  - Power Saving Mode
  - zPCR Information
- New z/OS Performance Support Overview
  - z/OS 1.13 Preview
  - z/OS 1.12
- Performance and Capacity Planning Topics
  - WLM
  - Other
- Addendum
  - Older APARs or Performance Information

www.ibm.com/support/techdocs

Introducing the zEnterprise

Bringing hybrid computing to a broader set of businesses

New!

IBM zEnterprise 114 (z114)

The next generation midrange mainframe delivering extensive growth options, flexibility, efficiency and improved price performance.

zEnterprise Unified Resource Manager

Centralized management of heterogeneous resources for simplification and resiliency

zEnterprise BladeCenter Extension (zBX)

Integrated IBM POWER7® blades, IBM System x blades, and High-performance optimizers and appliances*

*** Statement of Direction**

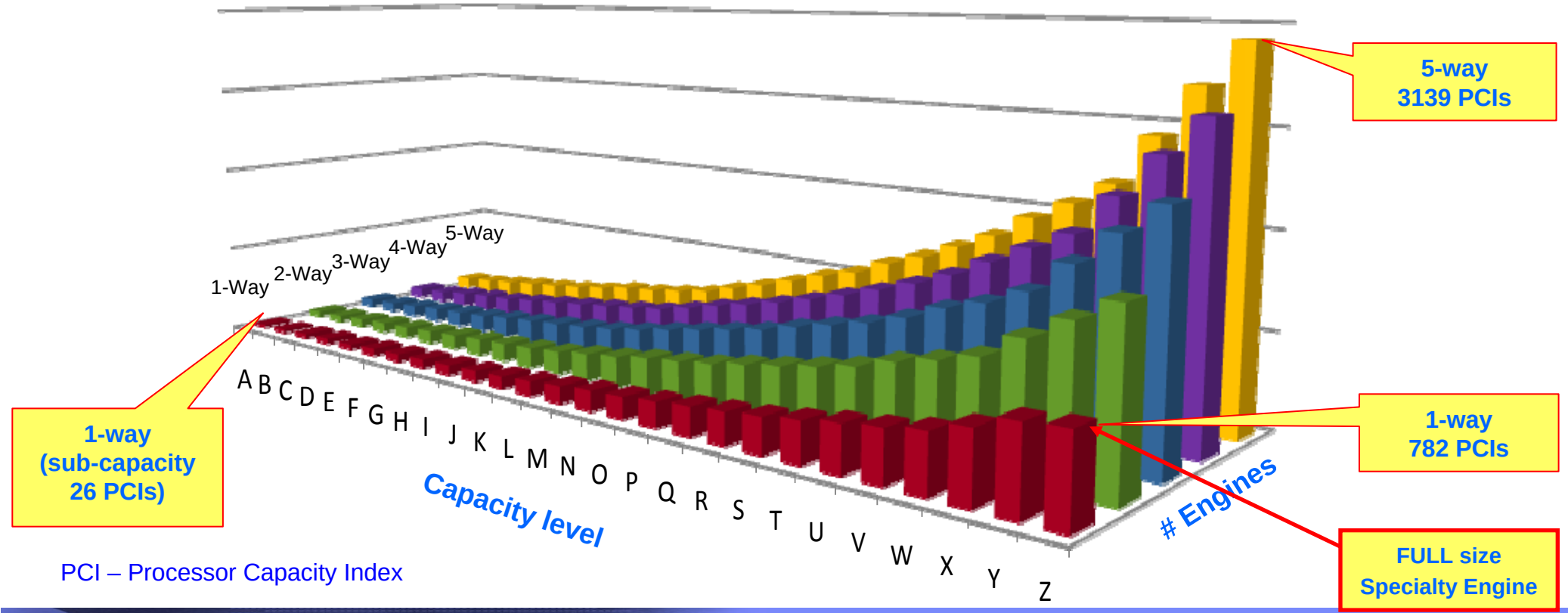
zEnterprise 114

- zEnterprise provides increased capacity in a single footprint
  - Designed for up to a 18% performance improvement per core and up to 12% improvement in total system capacity for z/OS, z/VM, and Linux workloads on System z compared to the z10 BC.
    - 12s0 technology
    - higher clock frequency 3.8 Ghz
    - out-of-order instruction processing
    - larger caches
    - compiler enhancements
- Connectivity improvements include bandwidth and throughput

z114 Sub-capacity Processor Granularity

- **The z114 has 26 CP capacity levels (26 x 5 = 130)**
  - Up to 5 CPs at any capacity level
    - All CPs must be the same capacity level
- **The one for one entitlement to purchase one zAAP and/or one zIIP for each CP purchased is the same for CPs of any speed.**
  - All specialty engines run at full speed
  - Processor Unit Value for IFL = 100

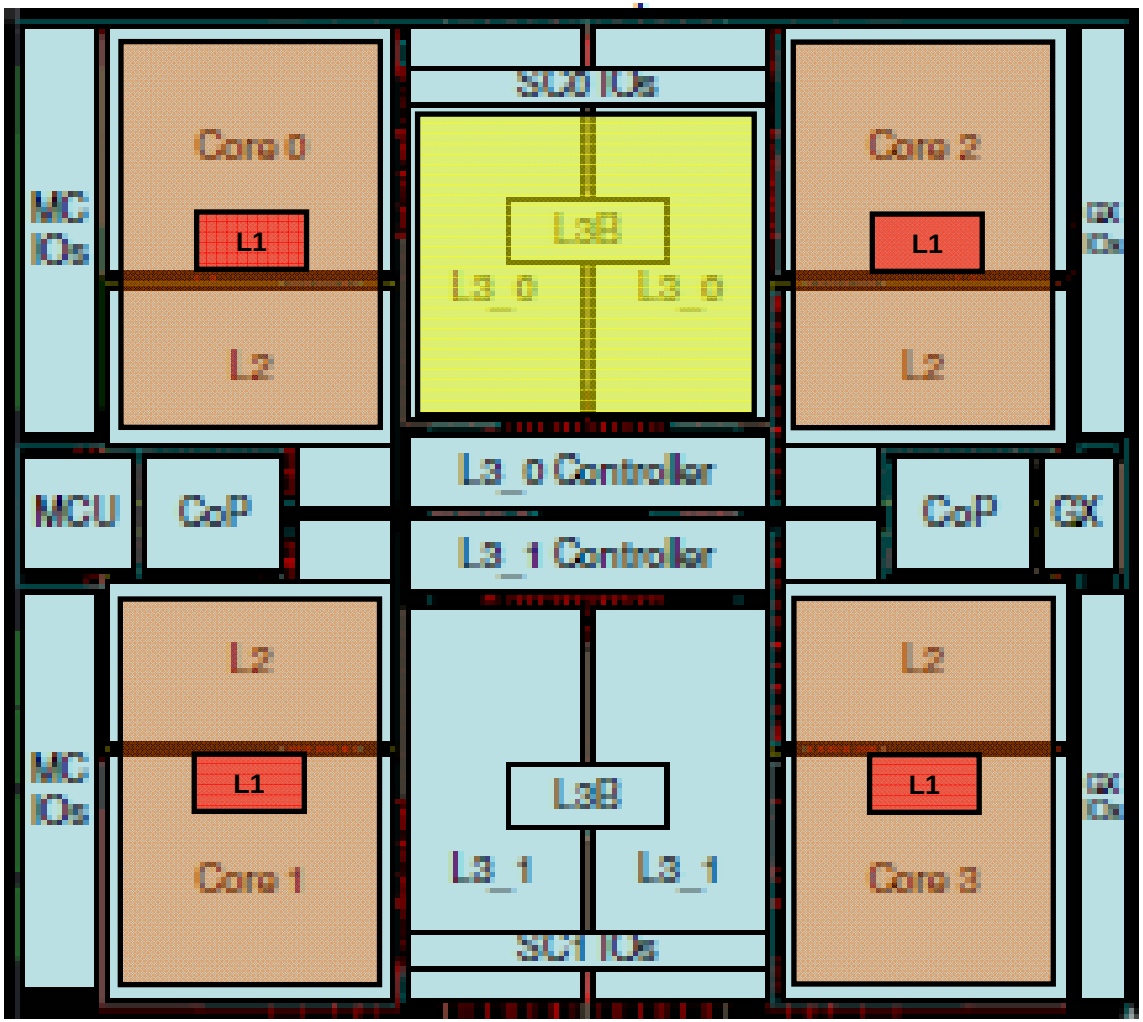
Number of z114 CPs	Base Ratio	Ratio z10 BC to z114
1 CP	z10 BC Z01	1.18
2 CPs	z10 BC Z02	1.16
3 CPs	z10 BC Z03	1.14
4 CPs	z10 BC Z04	1.13
5 CPs	z10 BC Z05	1.12



PCI – Processor Capacity Index

Single-Chip Module (SCM) in Processing Drawer(s)

- Quad core chips with 3 or 4 active cores
  - Same as the zEnterprise 196
- 3.8 GHz
- L1: 64K I / 128K D private/core
- L2: 1.5M I+D private/core
- L3: 12MB
  - Same chip as z196, but enabled half of the available 24MB
- L4: 96MB per processing drawer
  - On the SC Chip
  - 24MB assigned to each core
    - 24x4=96
  - Half of that on the z196



zEnterprise Information

- ***zEnterprise eXposed*** **- Introduction to zBX Performance Management and Monitoring**
  - **Part 1: The Intersection of WLM, RMF, and z/Manager Performance Management**
    Session: 10002, Tue. 11:00 AM
  - **Part 2: Experiences with the z/Manager Guest Platform Management Providers**
    Session: 10003, Tue. 1:30 PM
  - **Part 3: zManager and z/OS Workload Manager**
    Session: 10004, Tue. 3:00 PM

CPU Measurement Facility

- New hardware instrumentation facility available on z10 GA2, z196, and z114
  - New z/OS component - Hardware Instrumentation Facility (HIS)
  - Requires APARs OA25755, OA25750, and OA25773
  - Generates SMF 113.2 records
- Potential Future Uses:
  - Better workload characterization
  - ISV Product Improvements
  - Application Tuning
- Enhancements
  - z/VM support for CPU MF Counters via **APAR VM64961 (target: 8/19/2011)**
    - **z/VM 6.1 and z/VM 5.4 on z10s and z196s**
- New Support Information
  - **OA36816 - planned for August 12th, 2011**
    - Automates the process to allow HIS to continue running while indicating data loss in the SMF 113 records for the interval
  - Older but Important APARS
    - OA27623 - New fields added to SMF 113.2 records
    - OA30486 - Accept new parms on the MODIFY *hisproc* command
    - OA33052 - Support z196 extended counters

CPU Measurement Facility - Sampling

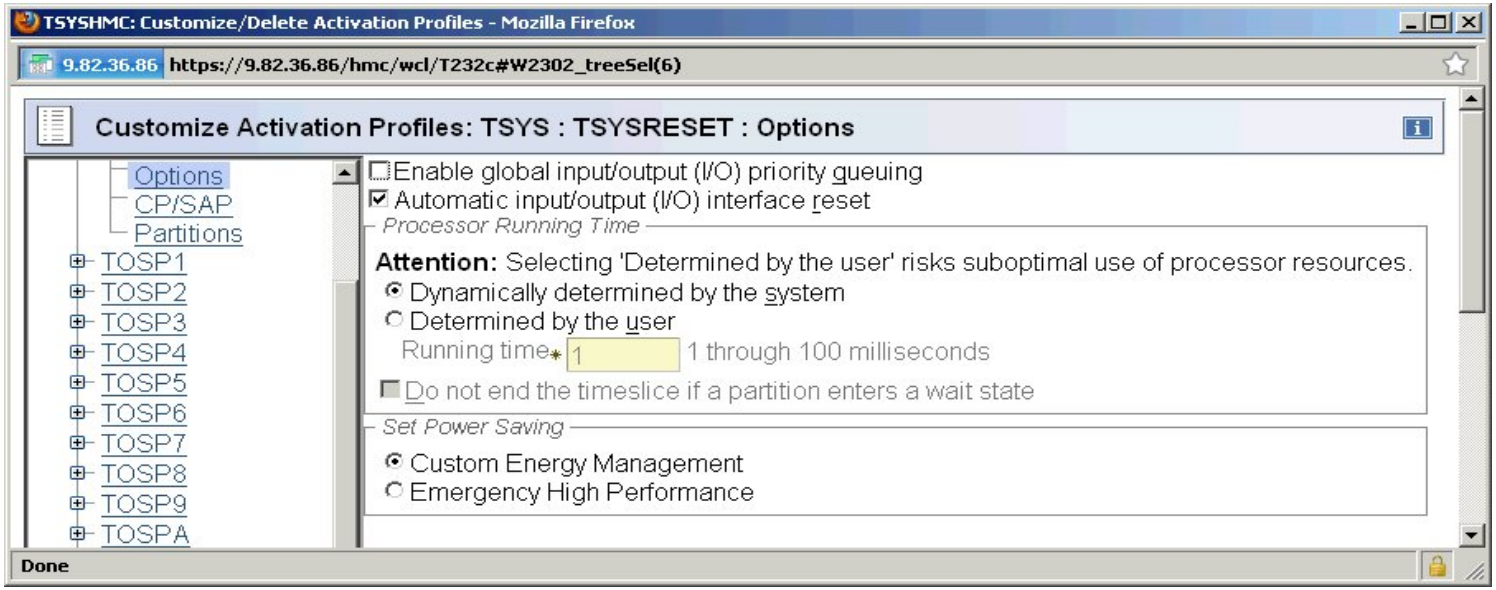
- Feb 2011 Hot Topics - A z/OS Newsletter - GA22-7501
  - "A whole lot of benefits from HIS data" article page 24
    - COUNTERS and an update on SAMPLING - HIS report tool and STG Lab Services
- Ensure the following PTFs are installed
  - z/OS Mapping
    - z/OS 1.9 APAR OA32113
    - z/OS 1.10 APAR OA32113 and APAR OA34485
    - z/OS 1.11 APAR OA30429 and OA34485
    - z/OS 1.12 APAR OA34485
  - CICS Mapping
    - APAR PM08568 (for CTS 3.2) or APAR PM08573 (for CTS 4.1)

CPU MF – 2011 Update and WSC Experiences
Session: 9999, Wed 3:00 PM

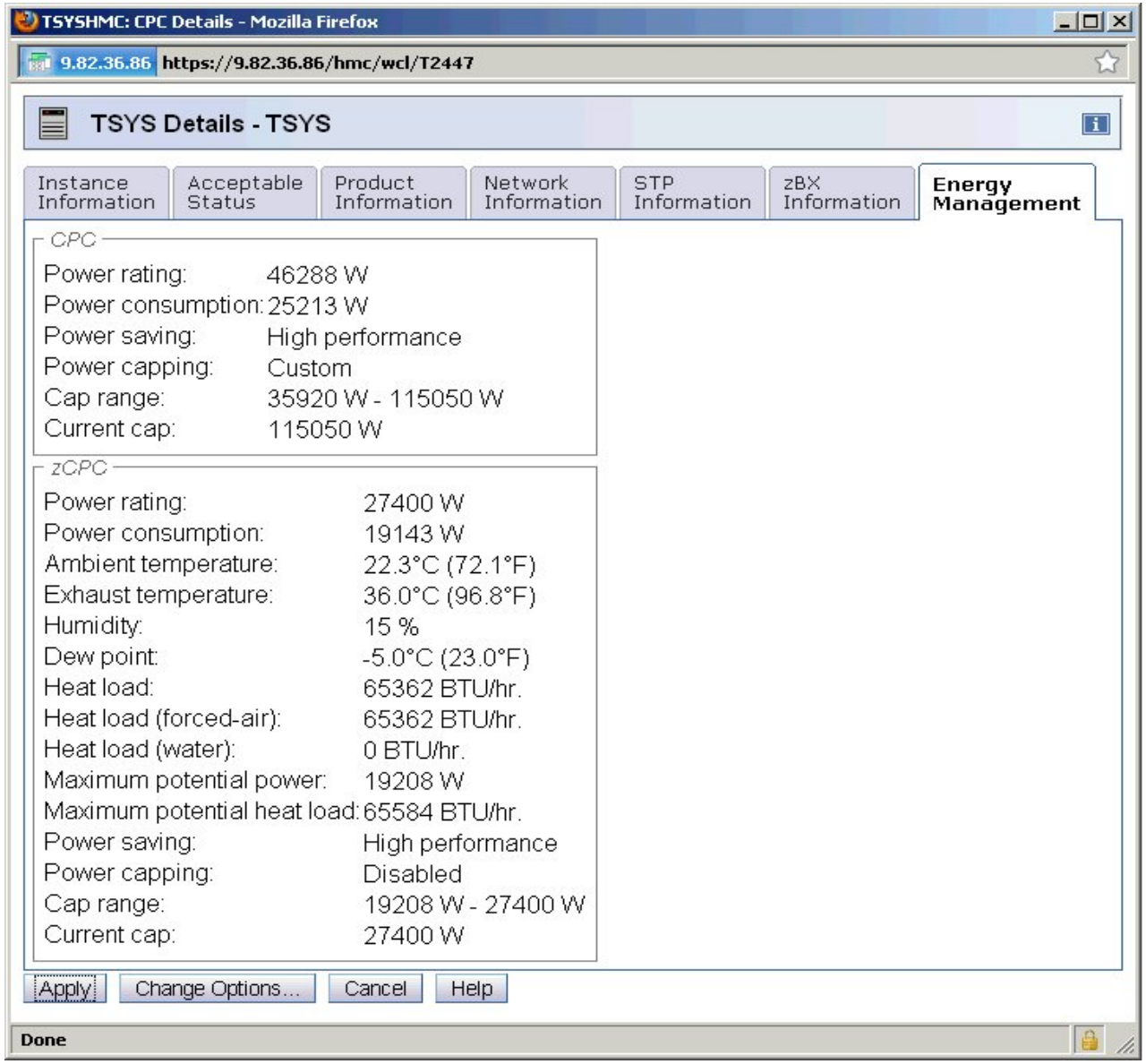
Exploring the SMF 113 Processor Cache Counters and LSPRs
Session: 9593, Thur 9:30 AM

z196 - Power Save Mode - Customer Initiated

- Reduce the energy consumption of your system
- Can be done on a scheduled basis
- A zCPC can be placed in power saving mode only once per day
- In z/OS when a Power Save event occurs:
  - SMF interval is ended and new one started
  - MSU and SU/SEC values are changed
  - SMF records record change (30, 70, 72, 89, 113.2, new 90.34)
  - Requires CPU times to be normalized, service units would be correct



z196 - Power Save Mode - Customer Initiated



z196 Power Save Mode

- Normal Power (Nominal)

```
                                                C P U  A C T I V I T Y

           z/OS V1R12                  SYSTEM ID SYSD             DATE 02/04/2011
                                       RPT VERSION V1R12 RMF      TIME 00.20.00
CPU        2817   CPC CAPACITY  6053          SEQUENCE CODE 00000000000C7675
MODEL      778    CHANGE REASON=NONE          HIPERDISPATCH=YES
H/W MODEL  M80



                                            W O R K L O A D   A C T I V I T Y
SYSTEMS
  ---ID---  OPT  SU/SEC  CAP% --TIME--  INTERVAL
  SYSD       00  58394.2  100 00.15.00  00.05.00
```

z196 Power Save Mode

- Power Save

```
                                                    C P U  A C T I V I T Y

           z/OS V1R12                SYSTEM ID SYSD              DATE 02/04/2011
                                     RPT VERSION V1R12 RMF       TIME 00.23.17
CPU         2817    CPC CAPACITY  5024          SEQUENCE CODE 000000000000C7675
MODEL       778     CHANGE REASON=POWERSAVE    HIPERDISPATCH=YES
H/W MODEL   M80



                                                W O R K L O A D   A C T I V I T Y
SYSTEMS
  ---ID---   OPT   SU/SEC   CAP% --TIME--   INTERVAL
  SYSD        00   58394.2    83 00.20.00   00.04.59
```

- CAP% - Percentage of effective capacity available to the processor
  - Value is 100 if the processor is working at its full, normal (nominal) capacity
  - If processor is working in power-save mode or cycle-steering mode, the value is less than 100

zPCR Latest Status

- Latest version is 7.4 (7/2011)
  - URL: http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS1381?OpenDocument&TableRow=4.1.0#4.1.
  - Includes z114 Support
- New Functions:
  - LPAR Configuration Capacity Planning function, Partition Detail Report window:
    - Test effect on capacity for the entire LPAR configuration with various alternative LCP count settings for shared GP partitions
      - Unparked LCPs only (as read from EDF or RMF)
      - Moderate or Minimum (based on partition weights)
      - User defined overrides
  - User's Guide and Online Help have been updated
  - QuickStart Guide has been updated
  - LSPR Document has been updated
  - Registration Improvements
    - When a CPS tool registration is required, for the benefit of those who are prevented from registering via the internet, the Register (e-mail) button is now immediately activated. Formerly, users had to attempt an internet registration before the Register (e-mail) button was activated

zPCR Latest Status

- Version 7.2a (1/2011)
  - Includes z196 Support
- New Functions:
  - New Support For Power Saving Mode
    - Adjustments to relative capacity values and MSU by clicking radio button
    - RMF/EDF input methods detect if data is in Power Save mode
  - HiperDispatch Changes
    - If reading partition info from EDF, for each z/OS partition represented with EDF, the number of logical CPs parked by HiperDispatch will be noted
  - Relative Nest Intensity Information
    - Workload selection will be enhanced via a "hint" if CPUMF data is provided via an EDF file
- Automated input available
  - RMF reports
  - EDF File
    - Requires an IBM supplied program (CP3KEXTR) to be run against SMF data of interest
    - Can read SMF 113 records to provide a "hint" regarding the relative nest intensity

z/OS 1.13 Performance Line Item Preview

- Response Time Distribution for Velocity Goals
- RMF GRS & Supervisor Delay Monitoring
- SMF IFASMFDL to stop reading before end of logstream
- zFS Direct I/O Support
- RMF Integrated Ensemble Performance Monitoring

z/OS V1.13 - Response Time Distribution for Velocity Goals

- Currently WLM reporting does not provide a response time distribution (ended transactions) for workloads with velocity goals

- Need to provide a response time distribution for all transactional workloads, even if they have a velocity goal
  - More data to analyze workload behavior and to detect problems
  - Better support for migration of goal definitions to response time goals

- IWMRCOLL to be updated to provide a response time distribution for service class periods with an execution velocity goal

- RMF Postprocessor Workload Activity report will displays the new response time distributions

Workload Management Update for z/OS 1.13 and 1.12
Session 10009: Mon. 4:30 PM

z/OS 1.13 RMF Report: Velocity R/T Distribution

```
REPORT BY: POLICY=POLICY01 WORKLOAD=STC SERVICE CLASS=STCDEF RESOURCE GROUP=*NONE PERIOD=1 IMPORTANCE=5
                                CRITICAL =NONE
-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---  --PROMOTED--  ----STORAGE----
AVG      28.04  ACTUAL           16.629  SSCHRT  89.0  IOC    524944   CPU    1.453  CP       0.22  BLK    0.000  AVG     1143.34
MPL      28.04  EXECUTION        15.724  RESP     0.2  CPU    649332   SRB    0.277  AAPCP    0.00  ENQ    0.000  TOTAL  32056.00
ENDED        2  QUEUED              904  CONN     0.1  MSO     14840   RCT    0.010  IIPCP    0.00  CRM    0.000  SHARED   200.56
END/S     0.00  R/S AFFIN             0  DISC     0.0  SRB    123890   IIT    0.197                LCK    0.000
#SWAPS     100  INELIGIBLE            0  Q+PEND   0.1  TOT      1313K  HST    0.000  AAP      0.00                -PAGE-IN RATES-
EXCTD        0  CONVERSION            0  IOSQ     0.0  /SEC     1459   AAP    0.000  IIP      0.00                SINGLE      0.0
AVG ENC   0.00  STD DEV               0                                 IIP    0.000                               BLOCK       0.0
REM ENC   0.00                                        ABSRPTN    52                                               SHARED      0.0
MS ENC    0.00                                        TRX SERV   52                                               HSP         0.0

GOAL: EXECUTION VELOCITY 20.0%     VELOCITY MIGRATION:   I/O MGMT  88.2%     INIT MGMT 88.2%

         RESPONSE TIME EX   PERF  AVG   --EXEC USING%--  -------------- EXEC DELAYS % -----------  -USING%-  --- DELAY % ---    %
SYSTEM                 VEL% INDX ADRSP  CPU AAP IIP I/O  TOT                                        CRY CNT  UNK IDL CRY CNT  QUI

*ALL        --N/A--    88.2  0.2  47.0  0.0 0.0 0.0 0.2  0.0                                        0.0 0.0   38  62 0.0 0.0  0.0
SYSD                   88.2  0.2  15.0  0.0 0.0 0.0 0.3  0.0                                        0.0 0.0   40  60 0.0 0.0  0.0
SYSE                   88.6  0.2  17.0  0.0 0.0 0.0 0.3  0.0                                        0.0 0.0   35  64 0.0 0.0  0.0

                                    ----------RESPONSE TIME DISTRIBUTIONS----------
SYSTEM: SYSD -----INTERVAL: 14.59.998 -----MRT CHANGES:  0 ---     SYSTEM: SYSE -----INTERVAL: 01.22.123 -----MRT CHANGES:  1 ---
   ----TIME----  -NUMBER OF TRANSACTIONS-  ------PERCENT------        ----TIME----  -NUMBER OF TRANSACTIONS-  ------PERCENT------
   HH.MM.SS.TTT  CUM TOTAL      IN BUCKET  CUM TOTAL IN BUCKET        HH.MM.SS.TTT  CUM TOTAL      IN BUCKET  CUM TOTAL IN BUCKET
<  00.00.00.200        581            581       94.2      94.2     <  00.00.00.300        581            581       94.2      94.2
<= 00.00.00.240        584              3       94.7       0.5     <= 00.00.00.360        584              3       94.7       0.5
<= 00.00.00.280        586              2       95.0       0.3     <= 00.00.00.420        586              2       95.0       0.3
<= 00.00.00.320        586              0       95.0       0.0     <= 00.00.00.480        586              0       95.0       0.0
<= 00.00.00.360        588              2       95.3       0.3     <= 00.00.00.640        588              2       95.3       0.3
<= 00.00.00.400        591              3       95.8       0.5     <= 00.00.00.600        591              3       95.8       0.5
<= 00.00.00.440        592              1       95.9       0.2     <= 00.00.00.660        592              1       95.9       0.2
<= 00.00.00.480        592              0       95.9       0.0     <= 00.00.00.720        592              0       95.9       0.0
<= 00.00.00.520        593              1       96.1       0.2     <= 00.00.00.780        593              1       96.1       0.2
<= 00.00.00.560        596              3       96.6       0.5     <= 00.00.00.840        596              3       96.6       0.5
<= 00.00.00.600        596              0       96.6       0.0     <= 00.00.00.900        596              0       96.6       0.0
<= 00.00.00.800        599              3       97.1       0.5     <= 00.00.01.200        599              3       97.1       0.5
<= 00.00.01.600        604              5       97.9       0.8     <= 00.00.02.400        604              5       97.9       0.8
>  00.00.01.600        617             13        100       2.1     >  00.00.02.400        617             13        100       2.1
```

z/OS 1.13 SMF Logstream Support

- Currently IFASMFDL will always read until the end of the logstream regardless of specified end date and time
- Two new options:
  - SMARTENDPOINT
    - First introduced with DUMP option in z/OS 1.12 with APAR OA31737 and OA34374
    - z/OS 1.13 extends support to ARCHIVE and DELETE
  - SMARTEPOVER(xxxx)
    - Specifies a value between 0000 and 0200 (2 hrs)
    - Default is 0200
    - SMARTEPOVER is added to SMARTENDPOINT to determine logical end point

z/OS 1.13 RMF GRS & Supervisor Delay Monitoring

- Collect and display system-wide contention information and contention information on address space level in
  - New SMF 72 subtype 5 record
  - New RMF XML Postprocessor Serialization Delay Report (SDELAY)
- New information
  - System Suspend lock types:
    - CMS
    - CMSEQDQ
    - CMSLatch
    - CMSSMF
    - Local
    - CML Lock Owner and
    - CML Lock Requestor
  - GRS lock types:
    - GRS Latch locks
    - GRS Enqueue Step
    - GRS Enqueue System and
    - GRS Enqueue Systems locks

RMF: The Latest and Greatest
Session 10011 – Monday 3:00 PM

Hiperdispatch White Paper V2

- Updated for the z196 and other common questions
- Discussion of meaning of MVS Busy with HD=YES
- Lists factors which influence potential HiperDispatch improvement
  - Processor cache technology
  - Number of physical processors
  - Size of the z/OS partition
  - Logical : Physical processor ratio
  - Memory reference patter
  - Exploitation of IRD Vary CPU Management
- Lists “Rule of Thumb” Expectations for z10 and z196
- Discusses importance of accurately set dispatch priorities for workloads

Hiperdispatch and LPAR

```
1                                              P A R T I T I O N   D A T A   R E P O R T
                                                                                                                              PAGE    3
             z/OS V1R10                SYSTEM ID LPAR1            DATE 04/29/2011            INTERVAL 14.59.998
                                CONVERTED TO z/OS V1R12 RMF       TIME 19.28.00              CYCLE 1.000 SECONDS
-
 MVS PARTITION NAME                  LPAR1          NUMBER OF PHYSICAL PROCESSORS          53                GROUP NAME        N/A
 IMAGE CAPACITY                         3165                  CP                           51                LIMIT             N/A
 NUMBER OF CONFIGURED PARTITIONS           4                  IIP                           2                AVAILABLE         N/A
 WAIT COMPLETION                          NO
 DISPATCH INTERVAL                   DYNAMIC
-
```

--------- PARTITION DATA		----MSU----			-CAPPING--		-- LOGICAL PARTITION PROCESSOR DATA -- PROCESSOR-		----DISPATCH TIME DATA----		-- AVERAGE PROCESSOR UTILIZATION PERCENTAGES -- LOGICAL PROCESSORS		--- PHYSICAL PROCESSORS ---		
NAME	S	WGT	DEF	ACT	DEF	WLM%	NUM	TYPE	EFFECTIVE	TOTAL	EFFECTIVE	TOTAL	LPAR MGMT	EFFECTIVE	TOTAL
LPAR1	A	494	0	582	NO	0.0	32.0	CP	02.17.24.319	02.20.44.154	28.63	29.32	0.44	17.96	18.40
LPAR2	A	446	0	762	NO	0.0	32.0	CP	03.01.28.607	03.04.05.167	37.81	38.35	0.34	23.72	24.06
LPAR3	A	59	0	0	NO	0.0	3.0	CP	00.00.00.000	00.00.00.000	0.00	0.00	0.00	0.00	0.00
LPAR5	A	1	0	0	NO	0.0	1.0	CP	00.00.00.000	00.00.00.000	0.00	0.00	0.00	0.00	0.00
PHYSICAL										00.10.58.833			1.44		1.44
									------------	------------			-----	-----	-----
TOTAL									05.18.52.927	05.35.48.155			2.21	41.68	43.90

Total LPAR weight = 1000

LPAR1 494/1000 = .494 * 53 CPs = 26.18 CPs

LPAR2 446/1000 = .446 * 53 CPs = 23.64 CPs

LPAR1 = 25 VH and 2 VM at 59% share (27 logicals unparked)

LPAR2 = 23 VH and 1 VM at 64% share (24 logicals unparked)

51 logicals unparked

53 physicals

Need to deactivate unused LPARs to reallocate their weight to VH and VM logicals

Hiperdispatch

- OA35989
  - On a large CEC with low utilization, except for a small test partition running with HD=YES, vertical low processors may not be unparked, even though there is sufficient demand on the small partition and there is a large amount of free capacity on the CEC
  - Routine which calculates free capacity suffered an overflow due to large amount of unused capacity
- OA35860
  - Running with HD=YES, vertical low processors may be unparked even though there is no unused capacity available on the CEC
  - WLM calculations of available capacity did not account for capacity used by *PHYSCAL partition
    - Impact is only when there is high Physical LPAR management time
- OA36459 - OPEN
  - Not calculating the capacity used by vertical mediums and vertical low processors correctly

HiperDispatch

- **OA36054**
  - Beginning with z/OS 1.13 when running on an IBM zEnterprise z196 the default for Hiperdispatch will be YES

Share of the partition - assumes 1.5 logical to physical ratio	Number of Physical CPs + zIIPs + zAAPs			
	<=16	17-32	33-64	65-80
0 <= share in processors < 1.5	0%	0%	0%	0%
1.5 <= share in processors < 3	2-5%	3-6%	3-6%	3-6%
3 <= share in processors < 6	4-8%	5-9%	6-10%	6-10%
6 <= share in processors < 12	5-11%	7-13%	8-14%	8-16%
12 <= share in processors < 24	-	8-16%	10-18%	11-21%
24 <= share in processors < 48	-	-	11-21%	12-24%
48 <= share in processors <= 80	-	-	-	14-26%

- **OA30476**
  - LPARs with >64 logicals must run with Hiperdispatch=YES

Workload Promotion

- OA30068
  - PDSE hang can occur on various PDSE latches due to address space getting swapped out by WLM while holding PDSE resources
  - PDSE contention couldn't be resolved by blocked workload support since latch holder was swapped out
    - Would require the address space to be made non-swappable
  - PDSE latch processing is changed to add SYSEVENT ENQHOLD function to allow SRM to boost the service of the latch holder
    - Improves swapin recommendation value
- OA35373 OPEN
  - Unilaterally swapped address spaces are not being exchanged swapped for long periods of time. Max sawp out time is set to the minimum or OUCBOUTT or 30 mins
  - This is too long to leave work swapped

W O R K L O A D A C T I V I T Y

--PROMOTED--
BLK 0.000
ENQ 0.649
CRM 0.000
LCK 7.376

Check and understand why there are CPU times in any service classes

z/OS 1.12 Performance Items

- WLM Enhancements
  - WLM Managed Initiators will consider the impact of resource group maximums when starting initiators
    - SMF 99 records updated to show reason for not starting
  - Improve Discretionary Work Throughput
    - Run discretionary work for a longer period of time before dispatching other discretionary work, while still interrupting it after short periods for non-discretionary work

Better Batch: Exploiting New Functions to Improve Batch Processing
Session: 9998, Tue. 9:30 AM

- RMF Enhancements
  - RMF changed to be able to read SMF records directly from SMF log stream improving ability to run reports with current data
  - Include information in the CPU Activity Report about how many units of work are running or waiting for a processor (CP, zIIP, or zAAP)
    - Same information is added to SMF Type 70 records

z/OS 1.12 Enhanced Reporting of Work Units

- New in-ready distribution of work units provides a more detailed view of the CPU demand than the in-ready distribution of address spaces
- Number of work units is presented per processor type (CP, zAAP, zIIP)
- Data is added to the SMF 70 records

```
z/OS V1R12                    SYSTEM ID SYSD
                              RPT VERSION V1R12 RMF
SYSTEM ADDRESS SPACE AND WORK UNIT ANALYSIS
  ---------NUMBER OF ADDRESS SPACES---------
  QUEUE TYPES            MIN       MAX       AVG

  IN                      73        74      73.4
  IN READY                 6         9       8.8

  OUT READY                0         0       0.0
  OUT WAIT                 0         0       0.0

  LOGICAL OUT RDY          0         0       0.0
  LOGICAL OUT WAIT        24        25      24.6

  ADDRESS SPACE TYPES

  BATCH                   10        10      10.0
  STC                     85        85      85.0
  TSO                      1         1       1.0
  ASCH                     0         0       0.0
  OMVS                     2         2       2.0
---------NUMBER OF WORK UNITS-------------
CPU TYPES              MIN       MAX       AVG
CP                       5        60       9.3
AAP                      0         0       0.0
IIP                      0         2       0.6
```

z/OS 1.12 Performance Items

- Change in CPU reporting
  - Joblog messages IEF374I and IEF376I are replaced by IEF032I and IEF033I
  - Maximum number of minutes now displayed with IEF032I and IEF033I is 99999
  - Previous messages truncated any CPU time greater than 9999
    - Job used 12301 minutes the IEF374I and IEF376I messages displayed 2301

DB2 and zIIPs

- DB2 Parallelism and zIIPs
  - Controlled by a CPU threshold. Once the threshold is met all child tasks are zIIP eligible
  - Parents are not zIIP eligible
  - Parent and child CPU time contribute to the CPU threshold
  - Can see any kind of work, CICS, IMS, TSO, batch using zIIP resources

- PM30468
  - DB2 V10 now supports CPU used for prefetch and deferred write to run on a zIIP processor
    - Without this APAR the CPU time is reported under the DB2 MSTR address space
    - When enclave created for this purpose the home address space is DB2 MSTR
    - Changed to allow creation of the zIIP eligible enclave under a service task whose home address space is DB2 DBM1

DB2 and z/OS

- PM12256
  - DB2 changes the redirection amount for zIIP offload for SQL requests via DRDA over TCP/IP to 60%
  - Provides performance benefit by reducing processor switching overhead for eligible zIIP workloads
- PM28626 (DB2) and OA35146 (z/OS)
  - zIIP utilization levels can become more variable after PTFs for PM12256 applied
    - Most visible when DRDA apps create extended duration work threads in DB2 (held cursors)
  - Impacts also seen where zIIP processor speed differs from general processors
    - Performance of a single DRDA SQL statement can experience more variation from one execution to the next especially for longer running SQL statements
    - Requires z/OS APAR OA35146
  - Enclaves with associated control structures not established by DB2 can result in unauthorized processor utilization

XES

- OA35117
- XES is changing the method in which it counts CF subchannel busy
  - Remove sensitivity to processor speeds which may cause over-reporting
  - Updates the z/OS subchannel tuning algorithm for the change in counting path busy
- XES currently increments the count of path busy conditions for every path busy condition encountered on the particular subchannel
- XES is changing this count to reflect the number of CF operations which experienced 1 or more path busy conditions on the particular subchannel

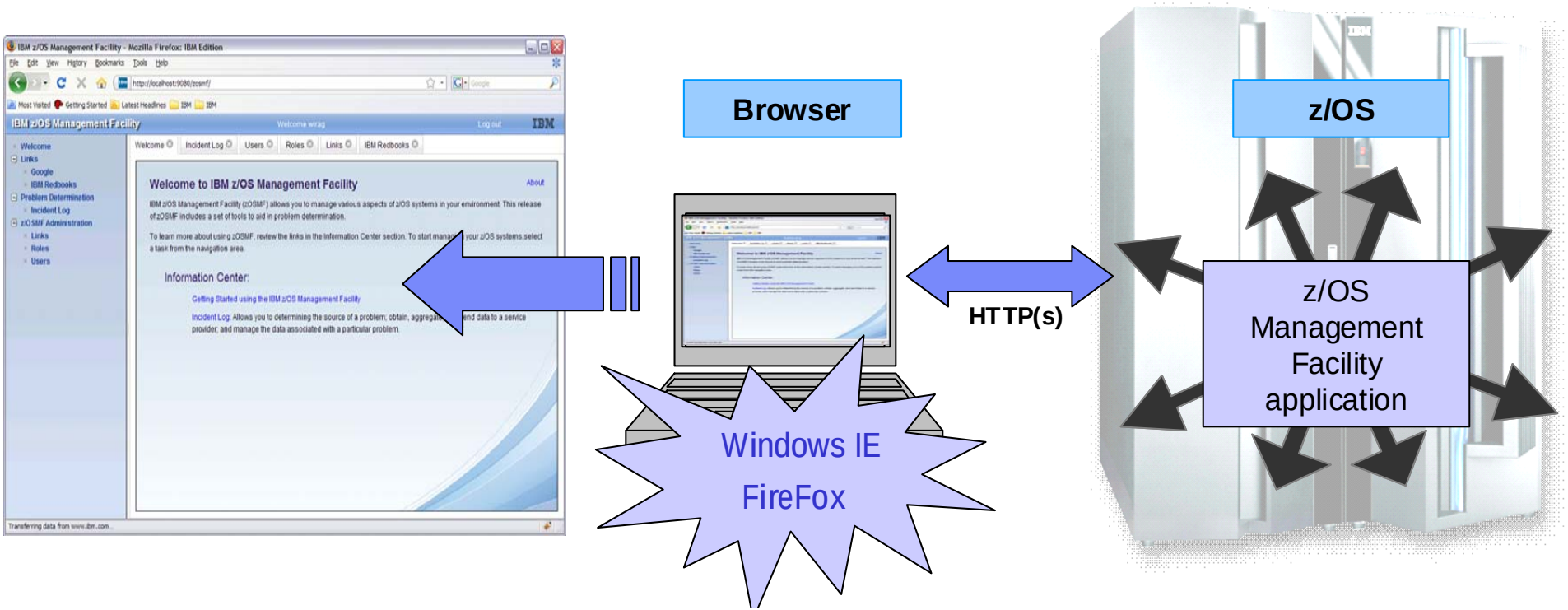
Migrating from z10 ICBs to z196 Infiniband- a Detailed Performance Study and User Experience
Session: A9743, Wed 9:30 AM

Websphere 6.1.0

- PM24445
  - IIOP, MDB, and internal work requests are showing inconsistent RMF Queue WAIT measurement
    - Some requests (HTTP) included the time waiting for the enclave to be selected by the worker thread in the servant
  - Introduce a new WAS directive to control calculation of queue time
    - wlm-enclave-exstartdefer=YES|NO
    - Default setting (YES) will include WLM Queue Wait time for all requests
  - In 6.1.0.39 Fix Pack

IBM z/OS Management Facility

Manages z/OS from z/OS



- z/OS Management Facility is an application on z/OS
  - Browser communicates with z/OSMF via secure connection, anywhere, anytime
  - Uses industry standards, such as Java™, DOJO, and CIM
  - Can exploit zIIP and zAAP engines, parts of z/OSMF use:
    - The z/OS CIM Server, Java
    - Workloads eligible for zAAP, or zIIP (with the zAAP on zIIP capability introduced with z/OS V1.11

IBM **z/OS** Management Facility

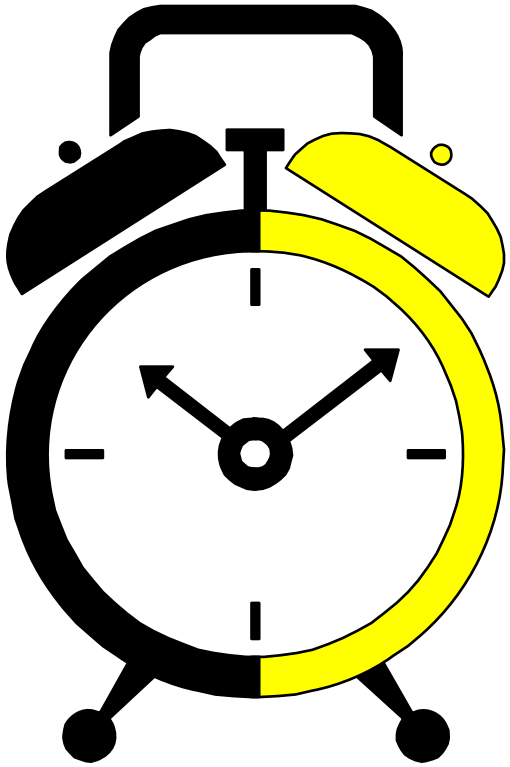
Welcome Page

- **Configuration**
  - **Configuration Assistant for z/OS Communication Server (R11)** – Simplified configuration and setup of TCP/IP policy -based networking functions
- **Links** to resources - provides common launch point for non -z/OSMF resources
- **Performance**
  - ★ **Capacity Provisioning** (R13*) - simplified monitoring of CP status for domains
  - **Resource Monitoring** (R12) – dynamic real time metrics for system performance
  - **System status** (R12) – single view of sysplex and Linux® performance
  - **Workload Management** (R12) – creation, editing, and activation of WLM policies
- **Problem Determination**
  - **Incident Log** (R11) – Simplified capture, packaging, and sending of SVC dump diagnostic data. (also avail with z/OS R10)
- **Software**
  - ★ **Deployment** (R13*) - Clone z/OS images and deploy software more easily and consistently, using a new z/OSMF software deployment task.
- **Storage**
  - ★ **DASD Management** (R13*) - Define new storage volumes to SMS quickly and easily using a single UI, using a new z/OSMF disk management task.
- **z/OSMF Administration** Authorization services, add users, define roles, add links.

Manage Your Workloads and Performance with z/OSMF
Session: 10012, Thur. 3:00 PM

Addendum

- Older flashes which should still be understood, or make you go Hmmmm.
- APARs which are still causing issues, even though they are old.



System zEnterprise 114 Functions and Features

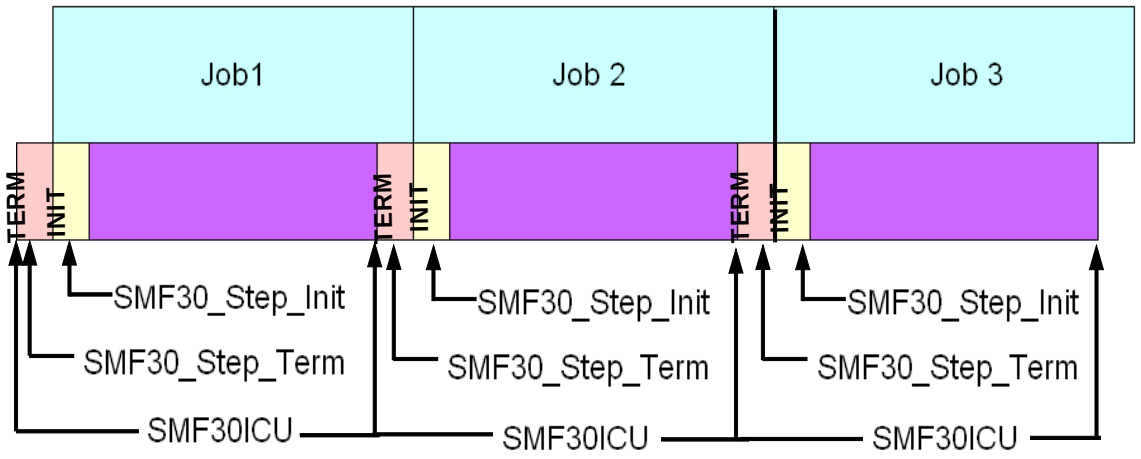
- Two hardware models
- Up to 10 processors configurable as CPs, zAAPs, zIIPs, IFLs, ICFs, or optional SAPs
- Up to 26 subcapacity settings across a maximum of 5 CPs
- Increased capacity processors
- Out of order instruction execution
- Improved processor cache design
- New and additional instructions
- Dedicated Spares on the Model M10
- Up to 248 GB of Redundant Array of Independent Memory (RAIM)
- Memory power save
- Cryptographic enhancements
- On Demand enhancements
- 6.0 GB/sec InfiniBand I/O interconnect

- 2 New OSA CHPIDs – OSX and OSM
- New 32 slot PCIe Based I/O Drawer
- Concurrent I/O drawer add, remove, replace
- Doubled HiperSockets to 32
- Physical Coupling Links increased to 72
- Doubled Coupling CHPIDs to 128
- CFCC Level 17 enhancements
- Optional High Voltage DC power
- Optional overhead I/O cable exit
- NRF Support with either top exit or bottom exit I/O and power.
- STP enhancements
- zBX Model 002 with ISAOPT, POWER7, DataPower and IBM System x Blades
- Platform Management from HMC

More Granularity and Greater Precision in CPU Timing

- SMF30ICU and SMF30ISB includes time:
  - Time spent in previous job's termination
  - Time spent during current job's step initialization

- New fields added to the CPU accounting section of the z/OS 1.12 SMF type 30:
  - SMF30ICU_STEP_INIT
  - SMF30ICU_STEP_TERM
  - SMF30ISB_STEP_INIT
  - SMF30ISB_STEP_TERM



New z/OS 1.12 Discretionary Batch Improvements

- TIMESLICES=1-255 (IEAOPTxx)

- Specifies number of timeslices a CPU-intensive address space or enclave with a discretionary goal should be given before a dispatchable unit of equal importance is dispatched

- Increasing this parameter might:
  - Increase processor delay for some CPU-intensive work
  - Decrease the number of context switches between equal priority work and therefore increase the throughput of the system

- Parameter only affects discretionary work that is CPU-intensive as determined by significant mean time to wait (MTTW)
  - As controlled by the CCCSIGUR parameter

- Default: 1

New z/OS 1.12 Discretionary Batch Enhancements

- CCCSIGUR=0-32767 (IEAOPTxx)
- Specifies the minimum mean-time-to-wait (MTTW) threshold value in milliseconds for heavy CPU users
  - Used to determine the range of MTTW values which are assigned to each of the ten MTTW dispatching priorities - x'C0' to x'C9'
  - Specified real time value is adjusted by relative processor speed to become SRM time to give consistent SRM control across various processors
  - Default Value: 45
- Used to differentiate Dispatch Priority of discretionary work
  - Work clumps at x'C9'
    - Appears all address spaces have short MTTW
    - CCCSIGUR is too large and should be decreased
  - Work clumps at x'C0'
    - Appears all work has large MTTW
    - CCCSIGUR is too small and should be increased
- Recommendation: start by doubling or halving the value

z/OS 1.12 Performance Items

- Shutdown and Restart Improvements
  - Address spaces allocating large numbers of data sets (e.g. DB2, batch) should see substantial reductions in shutdown and restart times
  - Changing subsystem initialization from serial to parallel for initialization routines listed in IEFSSNxx and a new BEGINPARALLEL keyword
  - XCFIPL time improved when using very large sysplex couple data sets
- RAS Enhancements
  - Improve capture performance for SVC dumps with substantial amounts of data on Auxiliary Storage
    - Internal IBM laboratory tests reduced capture time by over 60%
  - SADUMP will better prioritize data capture for address spaces, and dump a number of system address spaces first irrespective of their ASID numbers
    - Capture data needed most to diagnose system problems with a partial dump
    - Allow specification of additional address spaces to be added to the predefined list using a new ADDSUMM option
    - z/OS Best Practices: Large Stand-Alone Dump Handling Version 2
      http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD103286

Understanding SMF Record Type 120, Subtype 9

WP101342

- WebSphere Application Server for z/OS Version 7 introduced SMF 120 subtype 9
  - Presents a unified picture of the server activity
  - Collects most of the data currently spread across the other SMF 120 subtypes plus contains new information
  - WebSphere creates one subtype 9 record for every request the server processes
    - External requests (application requests)
    - Internal requests, such as when the controller "talks to" the servant regions
- Currently existing SMF 120 subtypes are continued and remain unchanged
- Paper discuss the structure and content, its related operational issues, and explain how you can make the best use of it