



IBM Advanced Technical Support

# z/OS Performance **"Hot"** Topics Session: 7772

## Share, Winter 2011

Bradley Snyder  
IBM Corporation  
Advanced Technical Support  
Washington Systems Center

# Trademarks

IBM*	FICON*	System z*
IBM (logo)*	IMS	System z10
ibm.com*	Lotus*	Tivoli*
AIX*	POWER7	WebSphere*
BladeCenter*	ProtectTIER*	XIV*
DataPower*	RACF*	zEnterprise
CICS*	Rational*	z/OS*
DB2*	System Storage	z/VM*
DS4000*	System x*	

## 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 there from.

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.

InfiniBand is a trademark and service mark of the InfiniBand Trade Association.

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.

### 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

- New Processor
  - ▶ IBM<sup>®</sup> zEnterprise<sup>™</sup> 196 (z196)
  - ▶ zPCR Information
  - ▶ CPU Measurement Facility
  - ▶ Power Saving Mode
  - ▶ IBM Smart Analytics Optimizer
- New z/OS Performance Support Overview
  - ▶ z/OS 1.13 Preview
  - ▶ z/OS 1.12
- Performance and Capacity Planning Topics
  - ▶ ZAAPZIIP Information
  - ▶ GRS
  - ▶ z/OS and DB2
  - ▶ WLM Information
  - ▶ New OPT parameters
  - ▶ XES
- Addendum
  - ▶ Older APARs or Performance Information

# www.ibm.com/support/techdocs



United States [ change ]

Search

Home
Solutions ▾
Services ▾
Products ▾
Support & downloads ▾
My IBM ▾

Welcome Kathy Walsh [Not you?] [ IBM Sign in ]

**Techdocs Library**

Flashes

Presentations & tools

Technotes & tips

FAQs

White papers

Solution scenario profiles

Customer support plans

Sizings

Auxiliary Material

Search Techdocs

Techdocs feedback

**Related links**

- Redbook publications
- IBM Software Support Handbook

## Techdocs - the Technical Sales Library



This site provides access to the Technical Sales Support organization's technical information databases. It gives you access to the most current installation, planning and technical support information available from IBM pre-sales support, and is constantly updated. You can browse or search these databases by date, document number, product, platform, keywords, etc.

**New to Techdocs?** Take a look at our [detailed introduction](#), which describes the document categories available (those listed on the navigation area on the left side of this page).

Rather than browse these categories, as a convenience you may enter a search of the full **Techdocs** database, or of any category you wish, here:

Search: All of the Techdocs Library ▾  Allow word variants

for:   "Fuzzy" search

Hits: 50 ▾ Order by: relevance ▾

Include docs updated: any time ▾ [Help for Search](#)

**Also available:** our [Advanced search](#), where you can select documents based on various assigned document attributes.

**New to Techdocs?**

Is this your first visit to **Techdocs** (the Technical Sales Library)?

[→ Learn more](#)

**Returning to Techdocs?**

Looking for what's new in the **Techdocs Library**?

[→ Latest updates](#)

**Need Technical Support?**

Looking for support resources or other documents and tools?

[→ Support & downloads](#)

Smarter Systems for a Smarter Planet



## IBM zEnterprise System – Best in Class Systems and Software Technologies

*A system of systems that unifies IT for predictable service delivery*



### Unified management for a smarter system: **zEnterprise Unified Resource Manager**

- Part of the IBM System Director family, provides platform, hardware and workload management
- Unifies management of resources, extending IBM System z® qualities of service across the infrastructure

The world's fastest and most scalable system:  
**IBM zEnterprise™ 196 (z196)**

- Ideal for large scale data and transaction serving and mission critical applications
- Most efficient platform for Large-scale Linux® consolidation
- Leveraging a large portfolio of z/OS® and Linux on System z applications
- Capable of massive scale up, over 50 Billion Instructions per Second (BIPS)

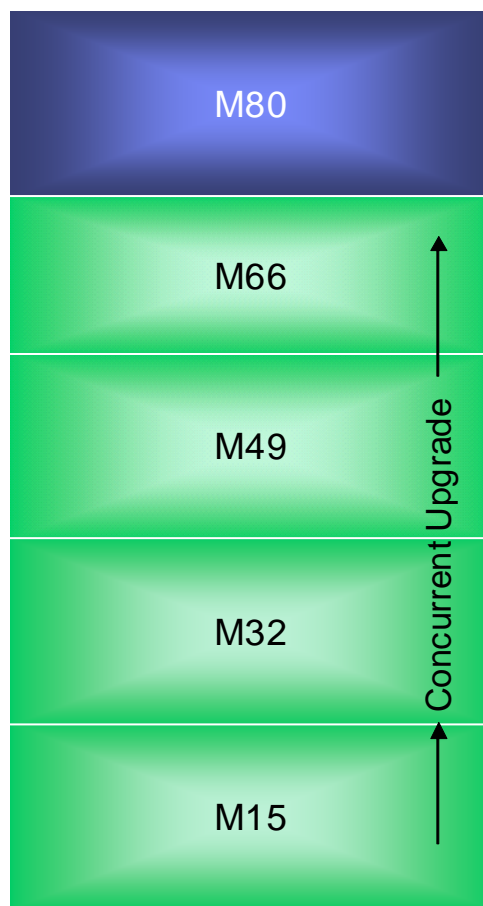
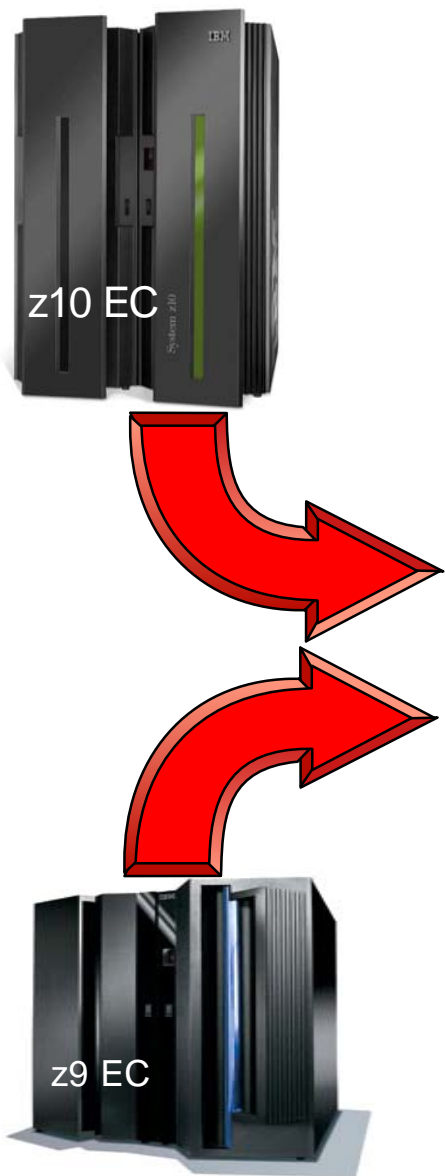
Scale out to a trillion instructions per second:  
**IBM zEnterprise BladeCenter® Extension (zBX)**

- Selected IBM POWER7® blades and IBM System x® Blades<sup>1</sup> for tens of thousands of AIX® and Linux applications
- High performance optimizers and appliances to accelerate time to insight and reduce cost
- Dedicated high performance private network



1 All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.

# System Offering Overview



**z196**

**z196 machine type: 2817**

**Processors:**

- ▶ 20 or 24 available cores per book
- ▶ Sub-capacity available up to 15 CPs
  - ▶ 3 sub-capacity points
  - ▶ 2 spares designated per system

**Memory for customer purchase:**

- ▶ System minimum = 32 GB
  - ▶ 16 GB separate HSA
- ▶ Maximum: 3TB / 768 GB per book
- ▶ Increments: 32 to 256 GB

**I/O Interconnects: (Same as z10)**

- ▶ 6 GB/sec
- ▶ Up to 16 per book (8 fanouts)
- ▶ Up to 48 per CEC (24 fanouts)

**zBX machine type: 2458**

# z196 versus z10 Hardware Comparison

## ■ z10 EC

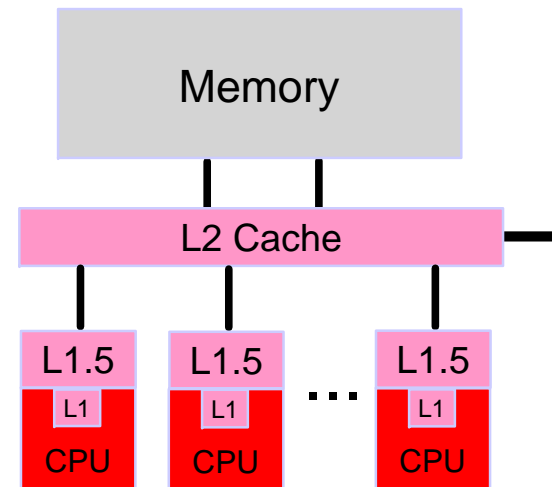
### ▶ CPU

- 4.4 GHz

### ▶ Caches

- L1 private 64k i, 128k d
- L1.5 private 3 MB
- L2 shared 48 MB / book

### ▶ Book interconnect: star



## ■ z196

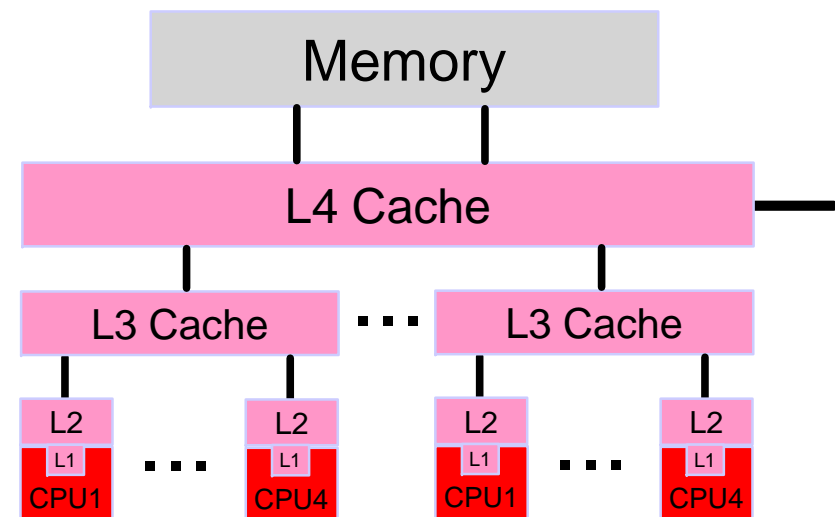
### ▶ CPU

- 5.2 GHz
- Out-Of-Order execution

### ▶ Caches

- L1 private 64k i, 128k d
- L2 private 1.5 MB
- L3 shared 24 MB / chip
- L4 shared 192 MB / book

### ■ Book interconnect: star



# CPU Measurement Facility

- New hardware instrumentation facility available on z10 GA2 and z196
  - ▶ 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
- **OA27623**
  - ▶ New fields added to SMF 113.2 records
    - Processor Speed for which the hardware counters are recorded, in terms of cycles/microsecond
    - New information on sampling frequency relative to processor speed
- **OA30486**
  - ▶ HIS has been enhanced to accept new parms on the MODIFY *hisproc* command
    - STATECHANGE - Can define action when significant event occurs during a HIS collection run
    - SMFINTVAL - Defines the interval SMF 113 records are recorded during a HIS collection run
- **OA33052**
  - ▶ HIS updated to support the additional counters provided in the z196 extended counter set

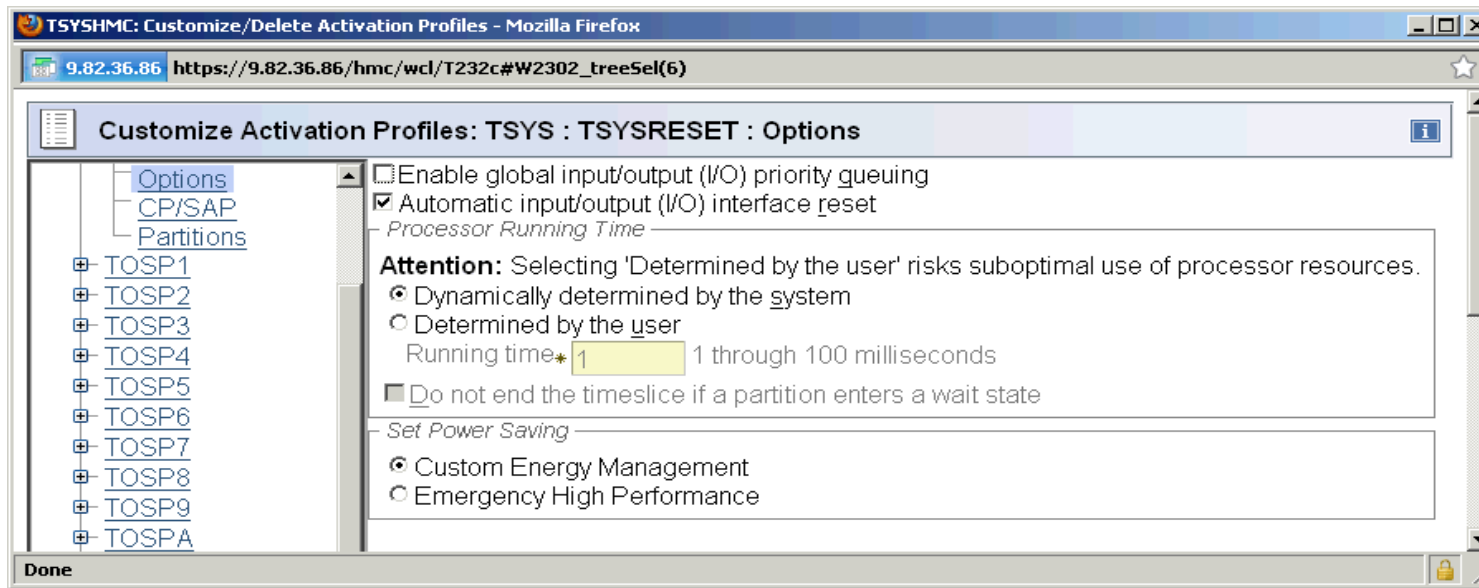


# CPU Measurement Facility - Sampling

- 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)

# 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

TSYSHMC: CPC Details - Mozilla Firefox

9.82.36.86 https://9.82.36.86/hmc/wcl/T2447

**TSYS Details - TSYS**

Instance Information | Acceptable Status | Product Information | Network Information | STP Information | zBX Information | **Energy Management**

**CPC**

Power rating:	46288 W
Power consumption:	25213 W
Power saving:	High performance
Power capping:	Custom
Cap range:	35920 W - 115050 W
Current cap:	115050 W

**zCPC**

Power rating:	27400 W
Power consumption:	19143 W
Ambient temperature:	22.3°C (72.1°F)
Exhaust temperature:	36.0°C (96.8°F)
Humidity:	15 %
Dew point:	-5.0°C (23.0°F)
Heat load:	65362 BTU/hr.
Heat load (forced-air):	65362 BTU/hr.
Heat load (water):	0 BTU/hr.
Maximum potential power:	19208 W
Maximum potential heat load:	65584 BTU/hr.
Power saving:	High performance
Power capping:	Disabled
Cap range:	19208 W - 27400 W
Current cap:	27400 W

Apply | Change Options... | Cancel | Help

Done

# 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 000000000000C7675
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 machine 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 7.2a (1/2011)
  - ▶ URL: <http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS1381?OpenDocument&TableRow=4.1.0#4.1>.
  - ▶ 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

## Related Sessions

- **Hardware Instrumentation Facility**
  - ▶ Exploring the SMF 113 Processor Cache Counters and LSPRs  
Session 8790, Thur. 9:30 AM
  - ▶ CPU MF - the “Lucky” SMF 113s - z196 Update and WSC Experiences  
Session 8882, Thur. 1:30 PM
  - ▶ Using And Getting Benefit From SMF 113 Records - Customer Experience  
Session 8531, Thur. 3:00 PM
  
- **Capacity Planning with LSPR**
  - ▶ To MIPS or Not To MIPS, That is the CP Question  
Session 8862, Thurs, 4:30 PM
  - ▶ The All New LSPR and z196  
Session 8864, Wed, 4:30 PM
  
- **zPCR Training**
  - ▶ zPCR Capacity Sizing - Part 1: Introduction and Overview  
Session 8883, Wed. 1:30 PM
  - ▶ zPCR Capacity Sizing Lab - Part 2 Hands On Lab  
Session 9098, Wed 3:00 PM
  - ▶ Usage Of zPCR Both In Performance Management And Capacity Planning Studies -  
Customer Experience  
Session 8532, Wed. 3:00 PM

# IBM Smart Analytics Optimizer

Capitalizing on the best of relational and the best of columnar databases

## What is it?

*The IBM Smart Analytics Optimizer is a workload optimized, appliance-like add-on that enables the integration of business insights into operational processes to drive winning strategies. It accelerates select queries, with unprecedented response times.*



## How is it different

- **Performance:** Unprecedented response times to enable 'train of thought' analyses frequently blocked by poor query performance.
- **Integration:** Connects to DB2 through deep integration, providing transparency to all applications.
- **Self-managed workloads:** queries are executed in the most efficient way
- **Transparency:** applications connected to DB2 are entirely unaware of ISAO
- **Simplified administration:** appliance-like hands-free operations, eliminating many database tuning tasks

**Breakthrough Technology Enabling New Opportunities**

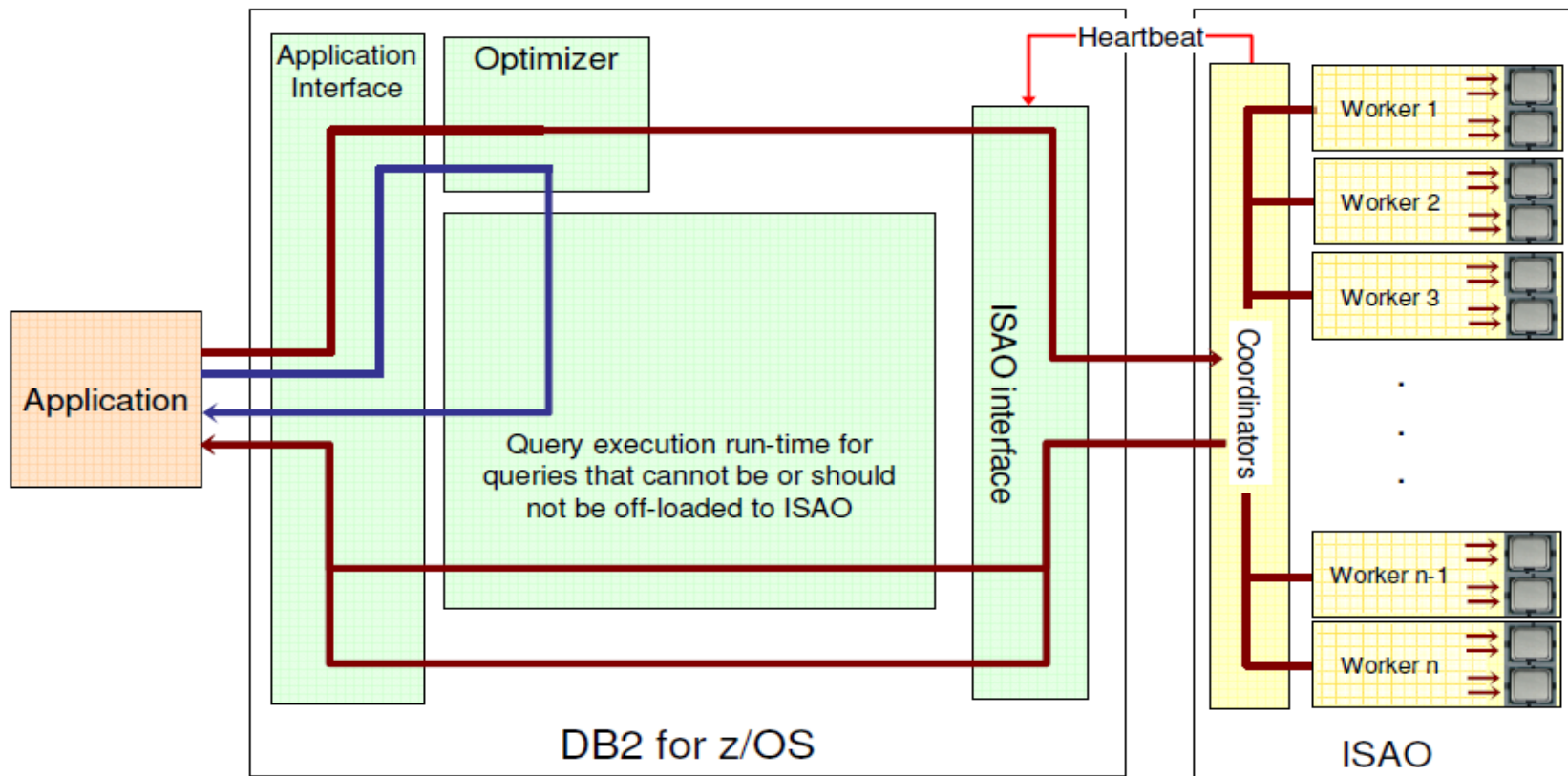
IBM Smart Analytics Optimizer ATS Beta Experiences on zBX model 002 and z196

Session: 8881 - Thur 11:00 AM



# IBM Smart Analytics Optimizer

**Total Solution Remains Centrally Managed by System z ...  
Without any Change to your Applications**



- Heartbeat (ISAO availability and performance indicators)
- Queries executed without ISAO
- Queries executed with ISAO

# WSC Benchmark Results

- Star schema structure database results

z196

Query	ET sec. on z	CPU sec. on z	ET secs. With Accel available	CPU secs. With Accel available	% Reduction in ET	Improvement Factor
WorkloadB_Q01	40.98	143.39	1.18	0.02	97.12%	34.73
WorkloadB_Q02	480.73	843.38	3.64	0.03	99.24%	132.07
WorkloadB_Q03	43.10	158.26	2.93	0.02	93.20%	14.71
WorkloadB_Q04	690.82	1312.67	2.38	0.02	99.66%	290.26
WorkloadB_Q05	1090.35	4.18	0.75	0.02	99.93%	1453.80
WorkloadB_Q06	449.28	757.59	2.46	0.02	99.45%	182.63
WorkloadB_Q07	328.49	784.10	2.47	0.02	99.25%	132.99
WorkloadB_Q08	357.94	824.28	2.63	0.02	99.27%	136.10
WorkloadB_Q09	474.47	1286.73	3.95	0.02	99.17%	120.12
WorkloadB_Q10	640.27	1858.90	3.89	0.02	99.39%	164.59
WorkloadB_Q11	883.42	65.09	2.29	0.02	99.74%	385.77
WorkloadB_Q12	573.04	1300.95	5.26	0.02	99.08%	108.94
WorkloadB_Q13	272.74	327.85	0.78	0.02	99.71%	349.67

# System z LPAR Considerations For Solution Edition

## ■ WP101841

- ▶ IBM Solution Edition is an IBM offering consisting of IBM hardware, software, and maintenance for customers choosing to run specifically identified workloads on IBM System z
- ▶ White Paper reviews the LPAR setup necessary to configure the System z hardware for Solution Editions
- ▶ Three z/OS environments where this SE can be enabled:
  - 1. Shared LPARs – No Previous Group Capacity
  - 2. Shared LPARs – Capacity Groups Currently Defined
  - 3. Dedicated z/OS LPAR Environment
- ▶ Though the paper addresses Solution Editions, many of the LPAR considerations and controls are applicable whenever adding new LPARs to an existing system

# 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

# IBM z/OS Management Facility

## Manages z/OS from z/OS

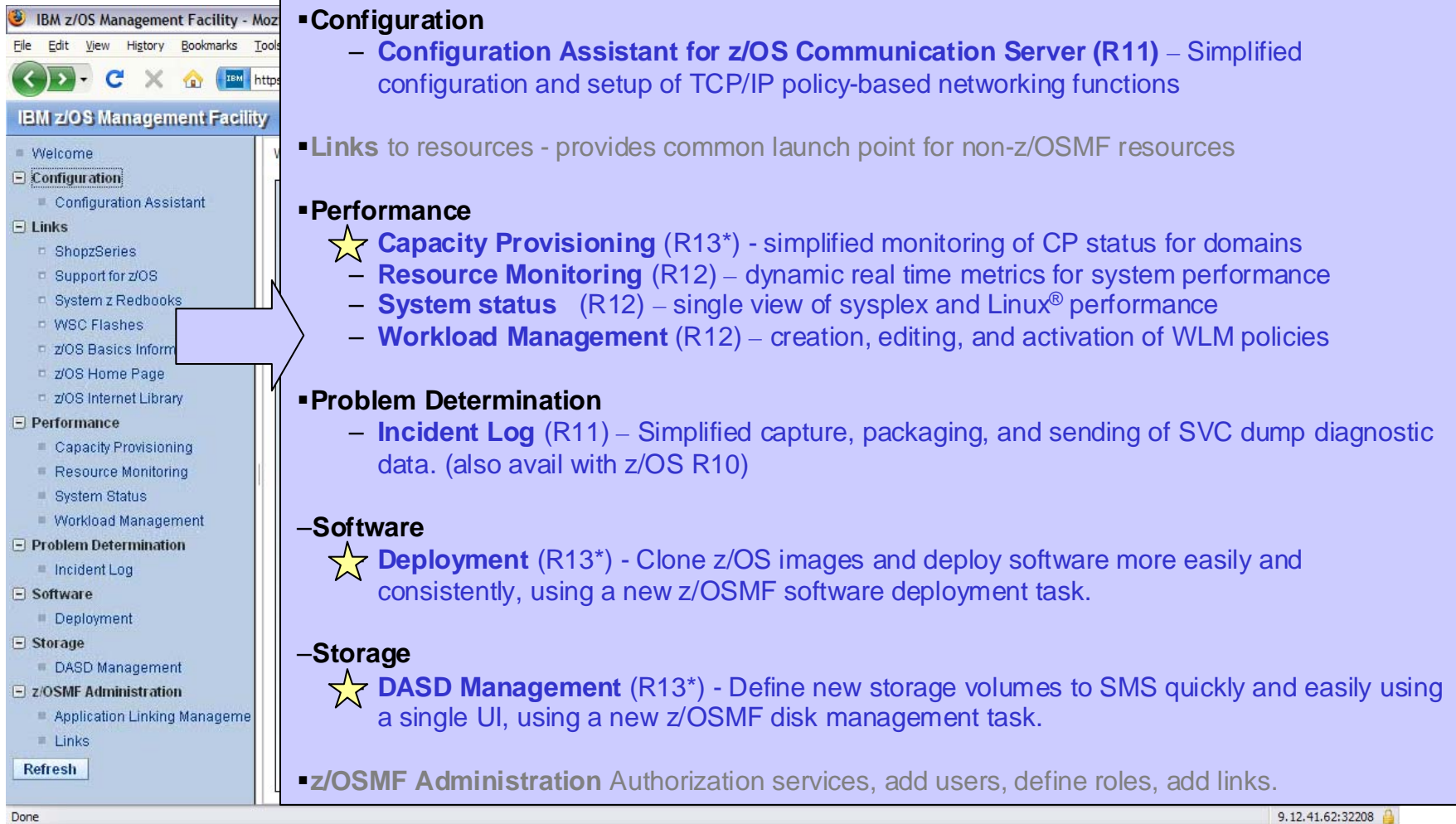


- **z/OS Management Facility is an application on z/OS**

- 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



The screenshot shows the IBM z/OS Management Facility web interface. On the left is a navigation menu with categories like Configuration, Links, Performance, Problem Determination, Software, Storage, and z/OSMF Administration. On the right is a list of features and services. A blue arrow points from the 'Performance' section in the menu to the 'Performance' section in the list.

- **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: 8859, Fri. - 11:00 AM

\*All statements regarding IBM's plans, directions, and intent are subject to change or withdrawal without notice.

## 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

### ■ 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 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/atmastr.nsf/WebIndex/TD103286>



## 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

## zAAPs on zIIPs

- When a [processor](#) has no zAAPs installed but the LPAR has zIIP(s), treat zAAP-eligible work instead as zIIP-eligible work
  - ▶ Exception: Under z/VM can define a guest without zAAPs, on a processor with zAAPs, to allow testing
- [IEASYSxx](#) system parm indicates whether zAAP on zIIP is enabled
  - ▶ ZZ=[NO](#) | YES for z/OS 1.9 and z/OS 1.10
  - ▶ ZAAPZIIP=[NO](#) | [YES](#) for z/OS 1.11 (also supports alias of ZZ)
  - ▶ The enablement state of zAAP on zIIP cannot be changed after IPL
- Timing fields within SMF will show offload time under zIIP
  - ▶ No method to determine what portion of zIIP time originated due to a zAAP request
  - ▶ Accounting and Capacity planning may need updating to use zIIP fields
- [APAR OA27495](#) for z/OS 1.9 and 1.10
- New operator command in z/OS 1.12 to display ZAAPZIIP eligibility  
D IPLINFO,ZAAPZIIP,STATE
- New Planning White Paper [WP101642](#):  
<http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP101642>
  - ▶ Must IPL into an environment which meets all of the criteria for ZAAPZIIP exploitation

# GRS Serialization Services - WP101725

- Reviews common causes of increased GRS CPU consumption and performance problems
  - ▶ Describes use of ENQ/DEQ Monitor
  - ▶ Operator command and Traces
  - ▶ Use of console Dumps to identify contention and hangs
  - ▶ Discusses Coupling Facility and XCF configuration with respect to GRS performance
  
- Case studies on how to approach and debug 5 common GRS Performance issues
  - ▶ TSO User hung
  - ▶ High GRS CPU Utilization
  - ▶ Lockout Cannot obtain GRS List lock
  - ▶ Global requests not being honored, system unresponsive
  - ▶ Contention Notification Request data queue grows very large
  
- **OA33898 - HIPER**
  - ▶ High CPU in GRS after migration to z/OS R1.11 processing GQSCAN signals from other systems in the sysplex
    - Increase CPU time (SRB) may be most noticeable on the GRS CNS
  - ▶ APAR applies to GRS only configs for serialization and to configs using other serialization products

## 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

### ■ OA32248

- ▶ DB2 V8 will specify `MANAGE_TASKS=YES` on WLM `IWMCONN` macro when creating DB2SPAS servants
  - WLM dynamically adjusts the number of TCB and address spaces as needed
  - Starts at 7 TCBs and adjusts as needed
- ▶ WLM incorrectly initializes a control block and additional TCBs may not be added

### ■ 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

### ■ JAVA, DB2, and ZAAPZIIP support

- ▶ Environment: Running JAVA on zIIPs, (ZAAPZIIP support), DB2 V8, type 2 connectors
- ▶ zIIP CPU time generated by the JAVA code is assigned to DB2 Class 1 time and not to ZIIP timing fields (Class 1 SE fields)
- ▶ Architecture of DB2 V8 is inadequate to correctly track ZAAPZIIP times
  - SMF 30 records are correct and can be used to adjust SMF 101 records
  - GCPU times are overstated in the SMF 101 records for DB2 V8
- ▶ Timing fields are correct for DB2 9

# 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

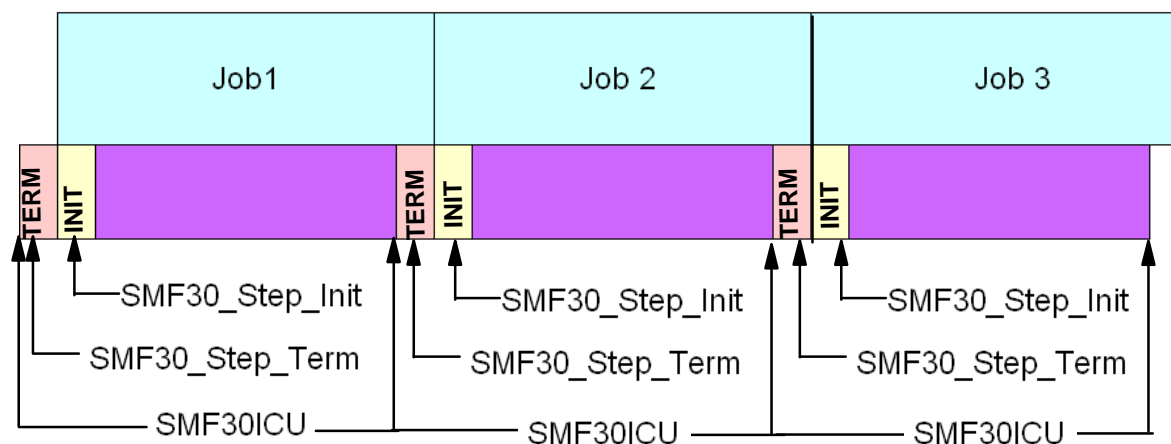
BATCH                                       10       10      10.0
STC                                        85       85      85.0
TSO                                         1        1       1.0
ASCH                                        0        0       0.0
OMVS                                        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

```

# 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

# 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

## Additional WLM Information

- **OA29314 - DOC - IRD and Group Capacity**
  - ▶ WLM only manages partitions in a Group Capacity which meet the following conditions:
    - Partition must not be defined with dedicated processors
    - Partition must run with Shared processors and Wait Complete=No must be set
    - Operating System must be z/OS 1.8 and above
    - z/OS cannot be running as a z/VM Guest
    - PR/SM Hard Capping is not allowed
  - ▶ Any LPAR not meeting the conditions is removed from the Group and the remaining LPARs are managed to the Group Limit
  - ▶ Group Capacity will function with IRD weight management as long as the partitions in the Group are not subject to capping
    - **No Weight moves will take place as long as the Group is being capped**

# RMF Enhancement - Operations

## ■ New RMF 1.11 Support to Display IEAOPTxx Active Member Settings

### RMF Monitor II Library List and OPT Settings

Enter selection number or command on selection line.

```

1 Link list          LNKLSTxx - Link Library list
2 LPA list          LPALSTxx - LPA Library List
3 APF list          IEAAPFxx - Authorized Program List

```

### 4 OPT IEAOPTxx - OPT Settings

```

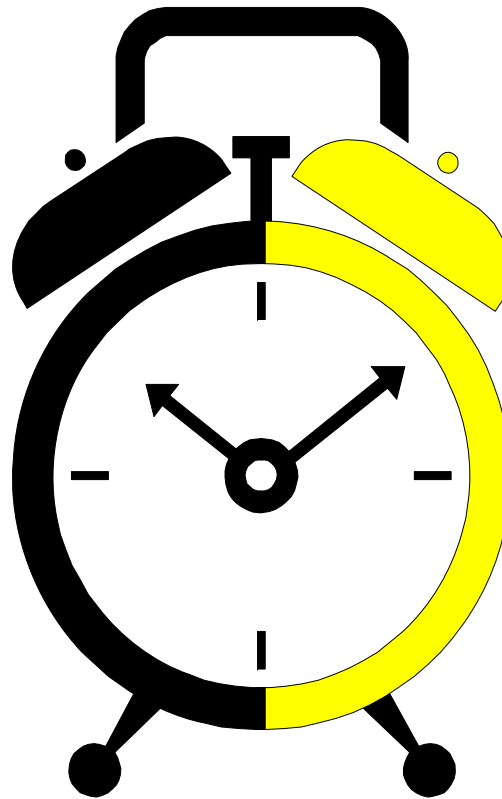
                                RMF - OPT Settings                                Line 1 of 29
                                CPU= 84/ 84 UIC= 65K PR=   0                                System= SYSD Total
OPT: KW                          Time: 07/31/10 15:17:47
-- Parameter -- - Default - -- Value -- Unit ----- Description -----
ABNORMALTERM                    Yes                Yes Y/N  Abnormal terminations in routing
BLWLINTHD                        20                 20 sec  Time blocked work waits for help
BLWLTRPCT                         5                  5 0/00  CPU cap. to promote blocked work
CCCAWMT                          12000              3200 usec  Alternate wait management time
ZAAPAWMT                          12000              3200 usec  AWM time value for zAAPs
ZIIPAWMT                          12000              3200 usec  AWM time value for zIIPs
CNTCLIST                          No                 No Y/N   Clist commands count individually
CPENABLE                          10,30|0,0          10,30 %   Threshold for TPI (low,high)
DVIO                               Yes                Yes Y/N  Directed VIO is active
ERV                             500                50000/F2 SU  Enqueue residency CPU Service/DP
HIPERDISPATCH                    No                 Yes/Yes Y/N  Hiperdispatch is desired/active
IFAHONORPRIORITY                  Yes                Yes Y/N   Allows CPs to help zAAPs
IIPHONORPRIORITY                  Yes                Yes Y/N   Allows CPs to help zIIPs
INITIMP                          0                  9/F2 #    INITIMP value/DP for initiators

```

**INITIMP=E!! in OPT**

# Addendum

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



# WLM Tools

<http://www-03.ibm.com/servers/eserver/zseries/zos/wlm/tools/wlmque.html>

## ■ WLM and the IBM z/OS Management Facility (z/OSMF)

- ▶ z/OSMF is a framework for managing various aspects of a z/OS environment through a Web browser interface
- ▶ WLM Support is added to this management environment
  - Provides support to create, edit and review WLM policies
  - Integrates many new functions to simplify the management of WLM policies, such as integrated best-practice checking during policy editing and a repository to store policies

## ■ WLMQUE

- ▶ ISPF based tool to display the application environments currently being used on your z/OS system
- ▶ View the number of started and active server address spaces, and the service classes being used as work queues for the application environments
- ▶ Tool can be used for any kind of application environment from WebSphere, DB2 or user specified types and applications

## z/OS 1.11 New Group Capacity Reporting

- CAPPING WLM% (percentage of time WLM capped the partition) is insufficient when the partition is member of a capacity group:
  - ▶ WLM% only tells to what extent a partition is **subject to capping** but not whether the partition was actually capped
  - ▶ WLM% is always 100 when the partition is capped with the phantom weight method although the partition does not use the capacity it is entitled to
  - ▶ WLM% is more a matter of how WLM caps the partition instead of how much it is being capped
  
- RMF does not report how much capacity is available for the group

G R O U P C A P A C I T Y R E P O R T										
GROUP-CAPACITY	PARTITION	SYSTEM	-- MSU --	WGT	-CAPPING--	- ENTITLEMENT -				
NAME	LIMIT		DEF	ACT	DEF	WLM%	MINIMUM	MAXIMUM		
-WSCCAP01	<u>636</u>	WSC11	MVS1	0	231	352	NO	100.0	224	636
		WSC12	MVS2	0	175	268	NO	100.0	170	636
		WSC13	MVS3	0	287	380	NO	100.0	242	636
-----			TOTAL		<u>693</u>	1000				

# RMF Group Capacity Enhancements

- The Postprocessor Group Capacity report displays the percentage of time where the partition was actually capped
  - ▶ Users of Capacity Groups can determine the available (unused) capacity for their group and whether the partition was actually capped:
  - ▶ SMF record 70 subtype 1 (CPU Activity) is extended
- The Postprocessor Partition Data report displays the available capacity for the group
- Greatly improved is the calculation of the remaining time until capping in Mon III CPC Capacity report

G R O U P   C A P A C I T Y   R E P O R T

GROUP-CAPACITY NAME	PARTITION LIMIT	SYSTEM	-- MSU -- DEF	ACT	WGT	-CAPPING-- DEF	WLM%	ACT%	- ENTITLEMENT - MINIMUM	MAXIMUM	
-WSCCAP01	636	WSC11	MVS1	0	225	352	NO	100.0	80.7	200	636
		WSC12	MVS2	0	175	268	NO	100.0	0.0	150	636
		WSC13	MVS3	0	237	380	NO	100.0	0.0	200	636
-----			TOTAL		637	1000					



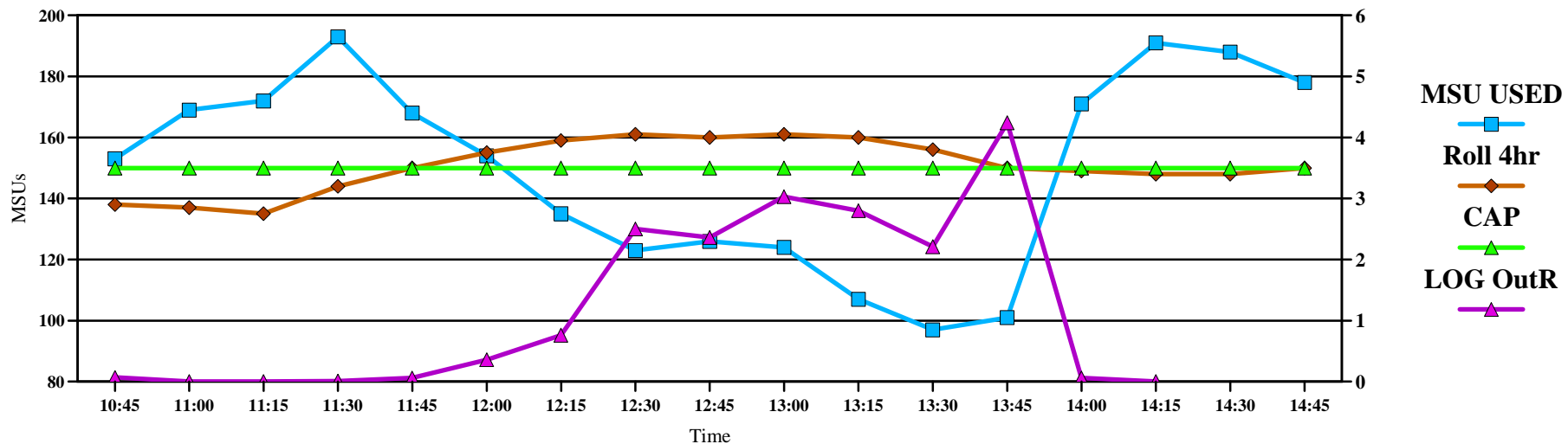
# Softcapping

## ■ OA28827

- ▶ z/OS 1.8 and up
- ▶ Unilaterally swapped address spaces in discretionary service classes are not being exchanged swapped in for long periods of time

## ■ OA32482

- ▶ z/OS 1.10 and 1.11
- ▶ When an LPAR is soft capped WLM will not start initiators in the resource adjustment path (underutilized system)



## z/OS 1.11 New Support for WLM Enclaves

- IWM4ECRE TYPE=WORKDEPENDENT
  - ▶ Resulting enclaves will be part of the creating TCB/SRB's transaction
  - ▶ Can now process parallel queries as a single transaction
  - ▶ If called while joined to an independent enclave a new type of enclave called work dependent is created:
    - Adopt owner address space and classification from the owning independent enclave
    - Extends an independent enclave's transaction
      - Work accumulates service as a group, and goes through period switch as a group
- Depending upon the environment the UOW is executing in when it calls IWM4ECRE with TYPE=WORKDEPENDENT the resulting enclave can be:
  1. Common dependent enclave
  2. Enclave of type work dependent
  3. Dependent enclave, which adopts owner address space and classification from the enclave out of which the service was called
- **APAR OA26104** for releases 1.8 and beyond

# TD104728 - z/OS Availability: Managing SVC Dumps to Mitigate Exhausting the Paging Subsystem

- Planning is required to ensure proper operation of SVC Dump
  - ▶ Too little AUX can cause a system outage or workload issues
- Discusses how to set MAXSPACE parameter
  - ▶ Shows how to review RMF Paging data to estimate how big a MAXSPACE your system can support
- Discusses how big a MAXSPACE parameter is needed
  - ▶ Impacts of Common and size of largest address spaces needing to be dumped
- Importance of ensuring a well configured AUX subsystem to support a large dump of multiple address spaces, or multiple dumps taken quickly in progression

# SVC Dump Planning

- As SVC Dumps grow in size it is important to properly define dump data sets
  - ▶ Especially as CHNGDUMP MAXSPACE parameters increase
- Do not specify multi-volume dump data sets as they are not supported by z/OS
  - Applies to both SYS1.DUMPxx and dynamically allocated data sets
  - Can cause SWA corruption in the DUMPSRV address space
  - Would only support a dump of 65535 TRKS, any dump requiring more than this would have a dynamic allocation failure

```
IGD17051I ALLOCATION FAILED FOR DATA SET LSE.SVC090.AMMLP1.D020321.T10
4951.S00002 , PRIMARY SPACE EXCEEDS 65535 TRKS
```

- Specify dump data sets as SMS managed, using a Data Class with the Extended Format attribute and use the Storage Class attribute Sustained Data Rate to get stripes
  - ▶ 3 volumes ("stripes") should be sufficient for a dump data set
- Additional Doc
  - ▶ FLASH10182 - z/OS Performance: SVC Dump Performance
  - ▶ DFSMS Storage Administration Reference, SC26-7402-09
  - ▶ DOC APAR - OA22623