



IBM Advanced Technical Support

z/OS Performance **"Hot"** Topics Session: 2100

Share, Summer 2010

Kathy Walsh
IBM Corporation
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[®] zEnterprise[™] 196 (z196)
 - ▶ zPCR Information
 - ▶ CPU Measurement Facility
- New z/OS Performance Support Overview
 - ▶ z/OS 1.11
 - ▶ z/OS 1.12
- Performance and Capacity Planning Topics
 - ▶ ZAAPZIIP Information
 - ▶ GRS
 - ▶ z/OS and DB2
 - ▶ WLM Information
 - ▶ Group Capacity
 - ▶ Hiperdispatch
 - ▶ ASM, SysTrace, Catalog, RRS, XES, JES2, etc.
- Addendum
 - ▶ Older APARs or Performance Information

www.ibm.com/support/techdocs

United States [change]

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

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: Allow word variants

for:

Hits: Order by: "Fuzzy" search

Include docs updated:

[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](#)

WSC

© IBM Corporation, 2010

4

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)

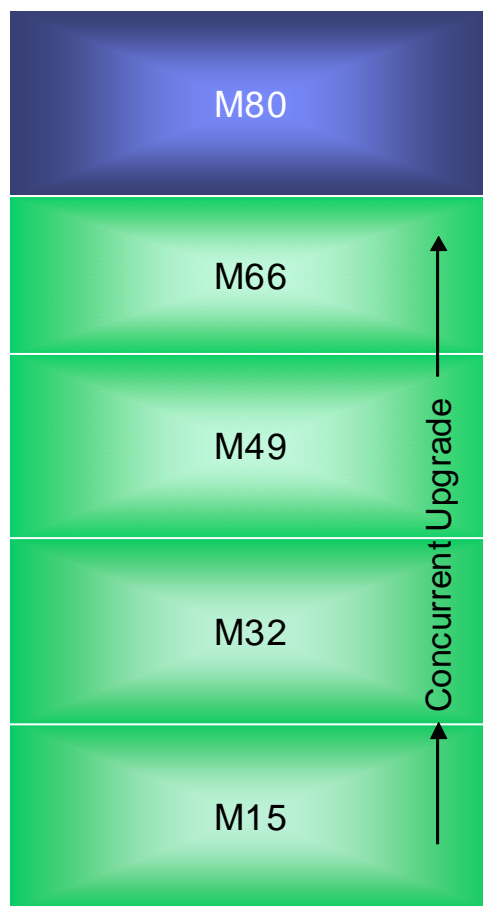
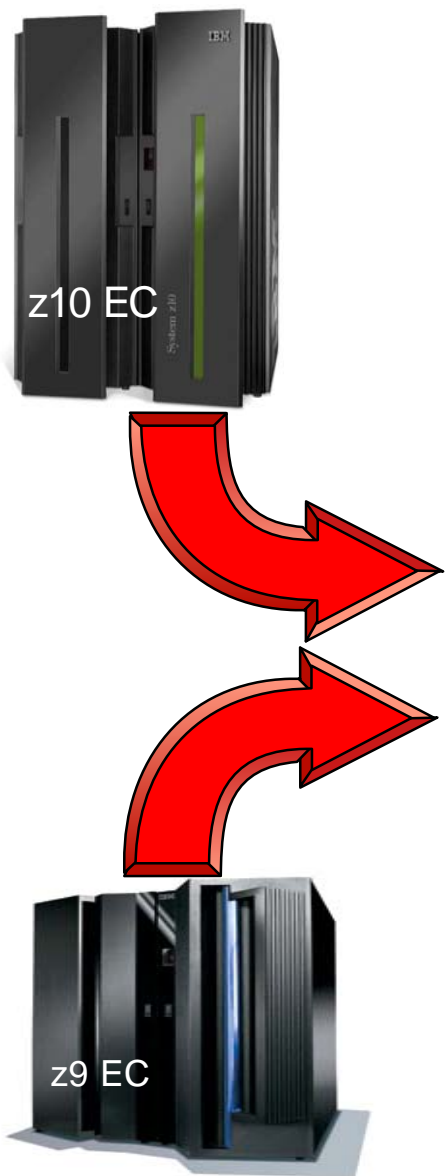
Scale out to a trillion instructions per second:
IBM zEnterprise BladeCenter® Extension (zBX)

- Selected IBM POWER7® blades and IBM System x® Blades¹ 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 represent 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

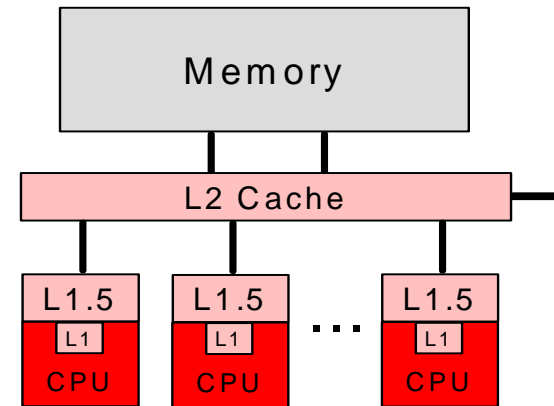
■ z10 EC

▶ 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

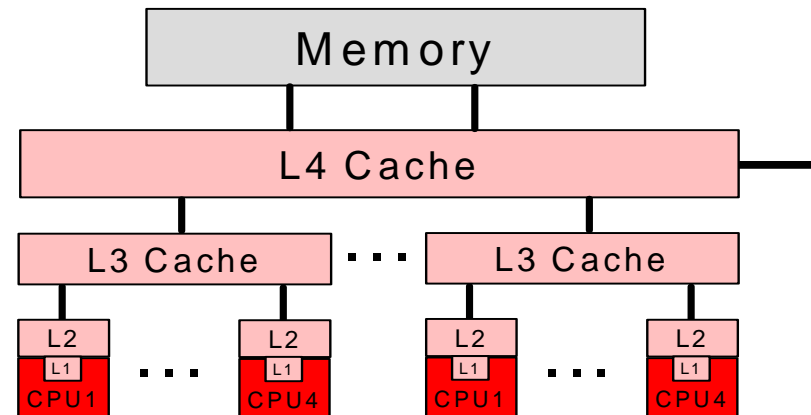
▶ CPU

— 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



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

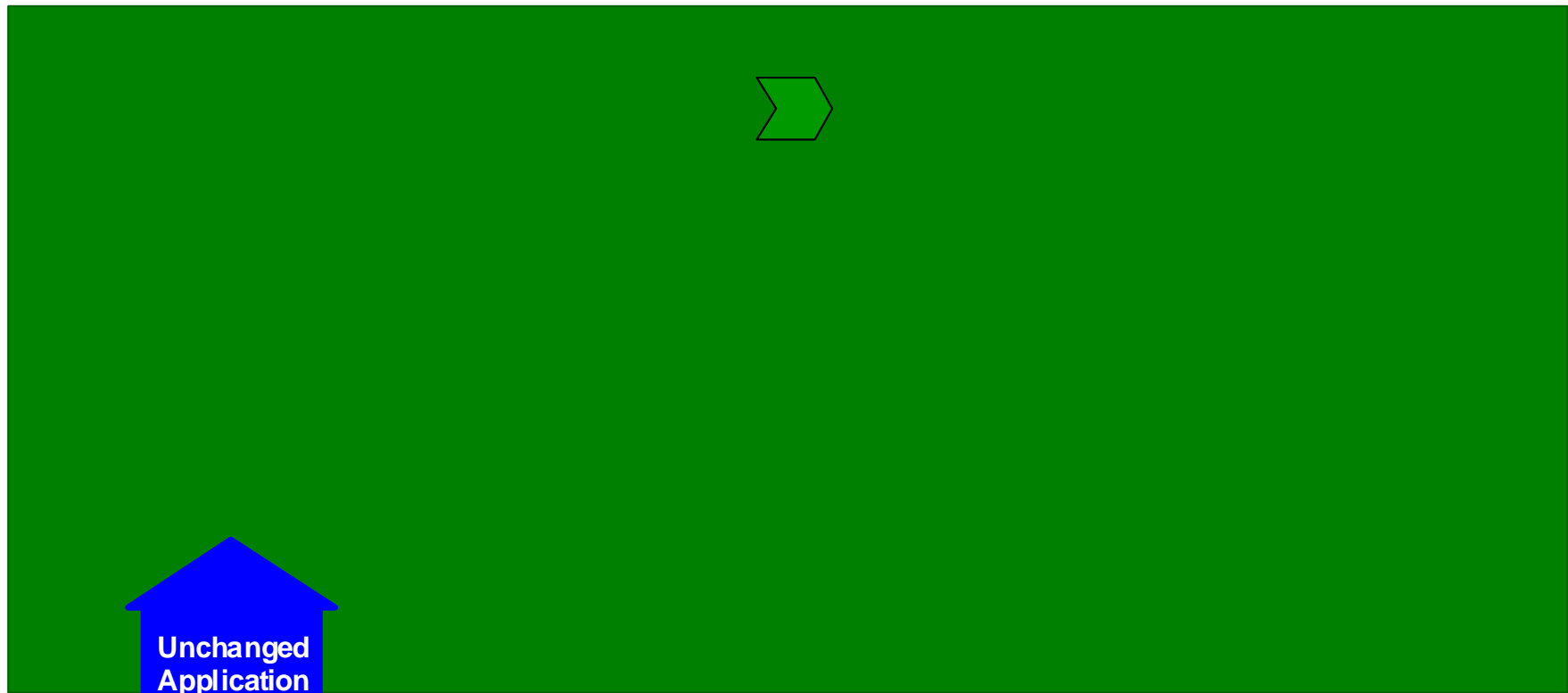
IBM Smart Analytics Optimizer Overview and WSC Beta Experiences, Thur, 3:00 PM

Accelerating your Queries using the IBM Smart Analytics Optimizer on System z, Tue, 9:30 AM

IBM Smart Analytics Optimizer

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

Total solution remains centrally managed by System z...



Unchanged
Application
Interface

...without any change to your applications.

zPCR Latest Status

- Latest version is 6.3c (5/19/2010)
 - ▶ URL: <http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS1381?OpenDocument&TableRow=4.1.0#4.1>.
 - ▶ z196 Support will be available middle of August (version 7.1)
- New Functions:
 - ▶ Reference-CPU controls have been restructured
 - A single (global) Reference-CPU window is defined
 - The Reference-CPU processor is limited to 1-way models
 - A Provisional Reference-CPU processor model can be temporarily assigned as any processor in the table when looking at LSPR Capacity Ratio Table
 - ▶ In Advanced-Mode, 4 LPAR configs can be defined and compared
 - ▶ LSPR Single-Image Processor Capacity Ratios table no longer available in Advanced-Mode
 - LSPR Multi-Image Processor Capacity Ratios table is preferred view for generalizing on processor model capacity
 - ▶ LSPR Multi-Image Processor Capacity Ratios table for z/OS-1.8 was removed
 - Only the table for the latest LSPR data will be provided (currently z/OS-1.9)
- Hands-On Training
 - ▶ [Lunch BOF: LSPR and CPU MF Update, Thurs, 12:15, Room 303](#)
 - ▶ [Framework for Doing Capacity Sizing on System z Processors, Thurs, 9:30](#)
 - ▶ [zPCR Capacity Sizing - Part 1: Introduction and Overview, Wed 1:30](#)
 - ▶ [zPCR Capacity Sizing Lab - Part 2 Hands On Lab, Wed 3:00](#)

z10 CPU Measurement Facility

- New hardware instrumentation facility available on z10 GA2
 - ▶ 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 MF - the "Lucky" 113s Update and IBM WSC Experiences, Wed, 9:30 AM

Exploring and Using the New SMF 113 Processor Cache Counters, Thurs, 9:30 AM

z/OS 1.11 Performance Items

- HiperDispatch improvements for zIIP processors
- New zAAP on zIIP support
- z/OS CIM server processing will be eligible to run on the System z zIIPs
 - ▶ Java based CIM client apps on z/OS can already use zAAPs
- New prefetch capability to provide perf improvements for XL C/C++ applications on z10
 - ▶ Reduce effects of memory latency by beginning to fetch data before it is known to be needed
- VSCR improvements in the base z/OS
 - ▶ GRS STAR mode sysplex-wide query requests
 - ▶ XML code page support
 - ▶ TCP/IP sockets processing
 - ▶ Allocation changed to significantly reduce the storage required by the Eligible Device Table (EDT) for many I/O configurations
- New SMF fields to record more information on use of Virtual, Real, and Aux storage above 2 GB

SMF23 Enhancements

- OA22414 - New Function - z/OS 1.8 and beyond
 - ▶ Supported on any IBM System z or IBM zSeries Processor
 - ▶ Adds new workload characterization counters to SMF 23 records
 - TCB and SRB dispatches per CPU
 - RSM Services intensity by tracking number of getmains, first reference page faults, non first reference faults, page fixes, etc.
 - Overall I/O rates
 - ▶ Better workload profiling data to enhance workload-mix selection for LSPR and zPCR

- z/OS 1.11 New Function
 - ▶ SMF 23 data is added to the SMF 70:71 records, and RMF Reports

P A G I N G A C T I V I T Y	
STORAGE REQUEST RATES	

GETMAIN REQ	415
FRAMES BACKED	148
FIX REQ < 2 GB	3
FRAMES < 2 GB	58
REF FAULTS 1ST	60
NON-1ST	0

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 nondiscretionary 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 planned to be 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>

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/atsmastr.nsf/WebIndex/WP101642>
 - ▶ Must IPL into a 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
 - ▶ Contention Notification Request data queue grows very large

DB2 parallelism, WLM, and zIIPs - History

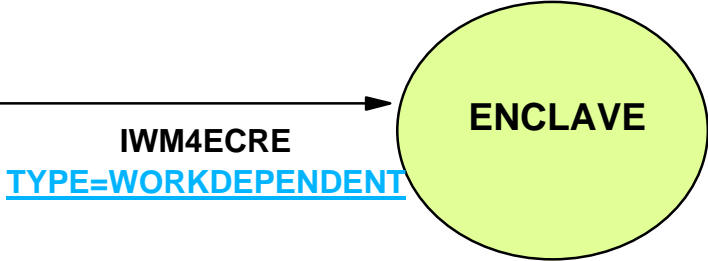
- DB2 Parallelism and zIIPs
 - ▶ Controlled by a CPU threshold. Once the threshold is met all child tasks are zIIP eligible
 - ▶ Parents are not zIIP eligible
 - ▶ Parent and child CPU time contribute to the CPU threshold
 - ▶ Can see any kind of work, CICS, IMS, TSO, batch using zIIP resources

- For SQL statements eligible for parallel execution DB2 may create new [Work Dependent Enclaves](#)
 - ▶ Without new Work Dependent Enclave support parallel enclaves must be classified using subsystem DB2
 - Enclave created in this manner are independent enclaves and are considered a new transaction
 - Must replicate existing classification rules for every subsystem which may run a stored procedure with the potential for exploiting CPU Parallelism

z/OS 1.11 New Support for WLM Enclaves

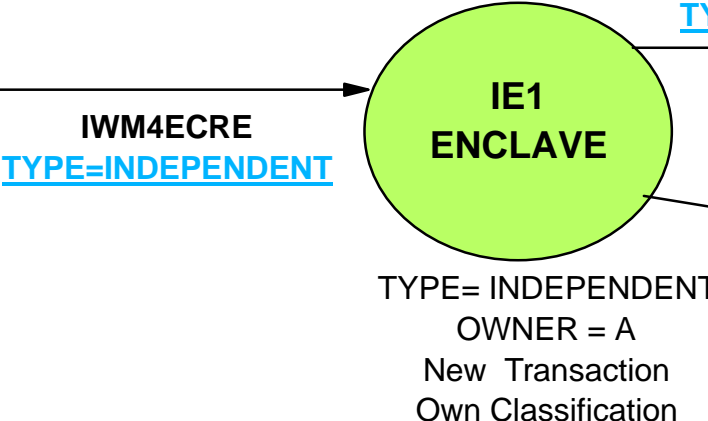
- **IWM4ECRE TYPE=WORKDEPENDENT**
 - ▶ Resulting enclaves will be part of the creating TCB/SRB's transaction
 - ▶ Can now process parallel queries as a single transaction
 - ▶ If called while joined to an independent enclave a new type of enclave called work dependent is created:
 - Adopt owner address 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

Addr Space A
TCB/SRB



TYPE= DEPENDENT
OWNER = A
Extended Transaction: A
Classification: A

Addr Space A
TCB/SRB



IWM4ECRE
TYPE=WORKDEPENDENT

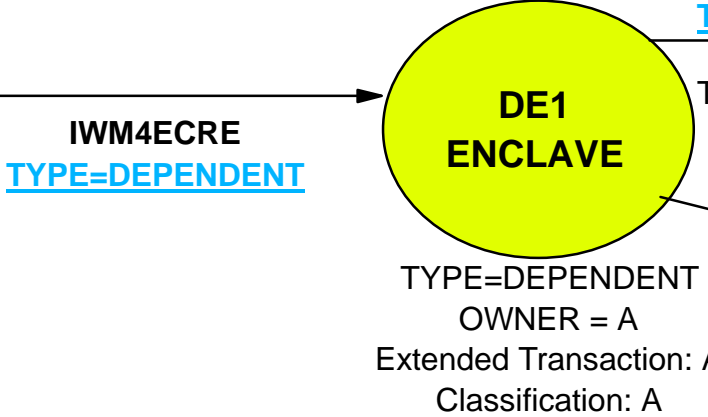
TCB/SRB running in any address space, joins IE1, calls IWM4ECRE with TYPE=WORKDEPENDENT



TYPE=WORK DEPENDENT
OWNER = A
Extended Transaction: IE1
Classification: IE1



Addr Space A
TCB/SRB

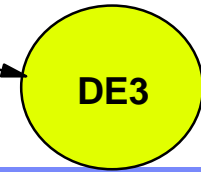


IWM4ECRE
TYPE=WORKDEPENDENT

TCB/SRB running in any address space, joins DE1, calls IWM4ECRE with TYPE=WORKDEPENDENT



TYPE=DEPENDENT
OWNER = A
Extended Transaction: A
Classification: A



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

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**

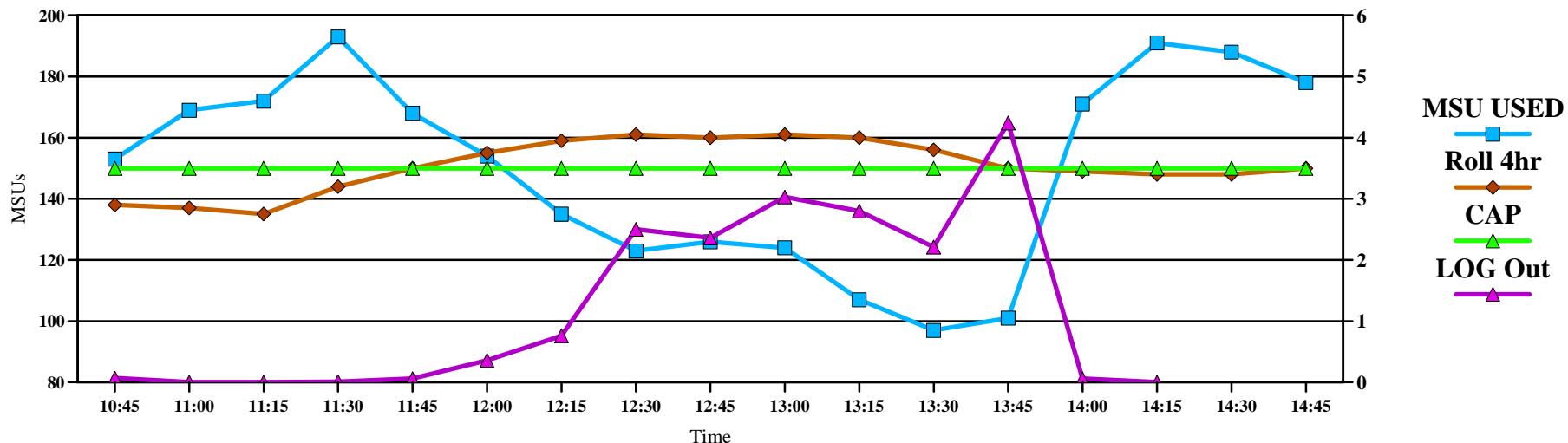
Softcapping

■ 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 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	PARTITION	SYSTEM	-- MSU --	WGT	-CAPPING--	- ENTITLEMENT -					
NAME	LIMIT		DEF	ACT	DEF	<u>WLM%</u>	<u>ACT%</u>	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		<u>637</u>	1000					

Session: RMF Latest and Greatest, Mon, 3:00 PM

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

Manage your Workloads and Performance with z/OSMF, Tue, 4:30

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

z/OS HiperDispatch

■ OA31733

- ▶ When all vertical low processors are parked, capacity of medium processors incorrectly set to zero when calculating available capacity of an affinity node
 - Results in capacity of the affinity nodes which these vertical medium processors belong to be estimated too low
 - As CEC approaches saturation, not enough workload is assigned to such affinity nodes and LPAR may under-utilize available capacity

■ OA32286

- ▶ z/OS R1.11 only with HD=YES
- ▶ High LPAR MGMT percentage and may also notice a lower capture ratio
- ▶ z/OS 1.11 was too aggressive in its SIGP signaling of idle CPs to process new work on a work unit queue

■ OA32383 and OA32081

- ▶ z/OS 1.9 and above, and HD=YES
- ▶ It is possible to see high PHYSICAL LPAR management time on a large n-way CEC with many LPARS
- ▶ z/OS was changed to control how aggressively processors are brought out of a wait

■ WSC Flash10526

- ▶ z/OS Performance: Capture Ratio Considerations for z/OS and IBM System z Processors V2
- ▶ <http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10526>

OA26789 - z/OS HiperDispatch

- z/OS 1.7 and beyond
- Running with HiperDispatch=YES a partition is running at or close to 100% logical busy and a vertical low processor may not be unparked even if the CEC is not 100% busy
- The range of the IEAOPTxx keywords ZAAPAWMT and ZIIPAWMT have been extended to 1600-499999 which corresponds to a timeframe from 1.6 to 500 ms
 - ▶ IEAOPTxx parms ZAAPAWMT and ZIIPAWMT control the wait time before a standard processor will help an assist processor
 - ▶ If HIPERDISPATCH=YES, the maximum value was 3200 which equates to 3.2 ms.
 - ▶ GCPs may help zIIPs or zAAPs earlier than desired
 - Value of 3200 may be too short and may result in GCP helping an assist processor instead of a vertical low zIIP or zAAP processor being unparked
 - Increased AAPCP/IIPCP time when zIIP/zAAP CPs are parked

OA32453 - ASM and PAVs (OPEN)

- z/OS 1.9 and beyond
- ASM paging operations relied upon the WLMPAV setting in HCD to be WLMPAV=YES to enable PAVs for paging operations
- HYPERPAV environments would most likely have set WLMPAV=NO since the keyword had no meaning in this environment
- Circumvention: Set WLMPAV=YES for all Page devices regardless of PAV type
 - ▶ Still important for releases prior to z/OS 1.9

System Trace - Real Storage Impact

- **II14465** - Performance Problems with System Trace
 - ▶ In z/OS 1.10 System Trace changed to allocate TTCH snaptrc buffers from 64bit storage
 - ▶ Minimum trace buffer size also increased from 256k to 1024k
 - ▶ On a system with a low amount of available 64 bit storage performance problems can be seen
 - ▶ Scenario:
 - System takes an ABEND and system trace would then request buffers above the bar
 - RSM then took a long time to search for appropriate frames
 - ▶ Circumvention:
 - Turn system trace off with the TRACE ST,OFF
 - Reconfigure the LPAR to just use 2GB of real, or add additional real storage to the LPAR (Provide more 64 bit available frames)
- **OA32947** (Open)
 - ▶ RSM searches excessively for a preferred high frame

z/OS 1.11 SRB Priority Adjustment

- Crippling performance problems can be seen if a flood of SRBs are scheduled from a higher priority address space to a lower priority address space
 - ▶ SRBs are given the dispatch priority of the target address space
 - ▶ Multiple XM POSTs can be the source of the SRBs
 - ▶ Existence of many SRB already dispatched in an address space may be due to CPU starvation

- New Support:
 - ▶ When there are many SRBs queued to be dispatched in an address space set the priority of new SRBs to the higher of:
 - The priority of the Scheduling address space
 - The priority of the Target address space

- [OA30154](#)
 - ▶ Provides SRB adjustment support to z/OS 1.9 and 1.10

JES2 Information

- **OA31416** - Hiper
 - ▶ Some Jobs may not be selected for execution due to failure to reregister service class
 - ▶ \$PXEQ then \$SXEQ will cause WLM service class re-registration

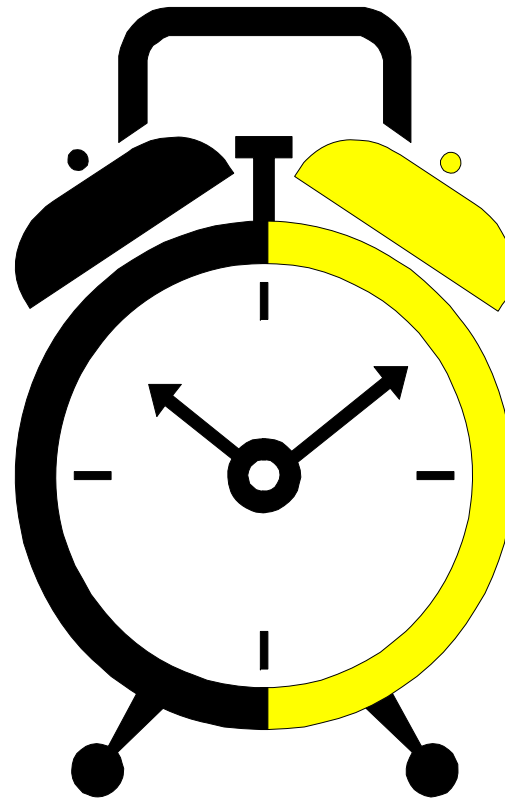
- **OA31814** - OPEN
 - ▶ WLM will not start any additional initiators
 - ▶ \$PXEQ then \$SXEQ to circumvent

- **OA31565** - Hiper
 - ▶ At z/OS 1.9 and above increased overhead using SDSF to invoke JES2 SSI for H, O, ST, I panels due to linkage stack inefficiencies
 - ▶ SDSF operations exhibit increased elapsed time and system sees drop in capture ratio
 - ▶ z/OS makes improvements in linkage stack management and JES2 takes advantage of improvement. Need PTF's for z/OS APAR [OA31897](#)

- **OA33063** - Hiper
 - ▶ Poor performance (high CPU) seen when using SDSF O panel searching held output
 - ▶ Changes storage management overhead and changes DESTID processing

Addendum

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



XES Information - OA28603

- Excessive sync-to-async conversion for coupling facility requests over short-distance ISC-3 coupling links is avoided via algorithm changes in the calculation of the sync-to-async conversion thresholds for CF requests
- New support to display the sync-to-async conversion thresholds currently in effect via the DISPLAY XCF,COUPLE command

D XCF,COUPLE

```
IXC357I 15.53.25 DISPLAY XCF 526
```

```
SIMPLEX SYNC/ASYNCHRESHOLD:                26
DUPLEX SYNC/ASYNCHRESHOLD:                  26
SIMPLEX LOCK SYNC/ASYNCHRESHOLD:            26
DUPLEX LOCK SYNC/ASYNCHRESHOLD:             28
```

RRS Information

- **OA29174** and **OA30351** - New Function
 - ▶ RRS Display command is being enhanced to aid in diagnosing the source of RRS transaction issues
 - The "comments" field on the existing console display command is expanded to include:
 - **A** - child or subordinate Unit of Recovery (UR) is waiting for application to signal it is complete and ready for syncpoint to begin
 - **E** - UR waiting for a Resource Manager to respond to a syncpoint exit
 - **P** - A Coordinator UR is waiting for a response from RRS on a remote subordinate system
 - Display RRS updated with new keywords:
 - **Family**: List of all URs involved in a cascaded transaction - either local or Sysplex
 - **URStatus**: Count of currently active RRS transactions by syncpoint state
 - **UREXceptions**: report of the URs currently WAITING

OA31115 - VSM and IMS New Function - HIPER

■ AbendU1011 on Restart of IMS

- ▶ Free blocks of ECSA have become sufficiently fragmented to prevent large getmain requests being satisfied
- ▶ For GETMAIN and STORAGE OBTAIN, in certain select requests for (E)CSA storage (using options STARTBDY or CONTBDY), z/OS performs a "first fit" within the free areas of storage, which can result in fragmentation of (E)CSA area
- ▶ A new **DIAGxx** option is provided which controls algorithms by which GETMAIN and STORAGE OBTAIN handle (E)CSA allocations

VSM BESTFITCSA(NO|YES)

NO: Use a "first fit" algorithm in certain situations such as when the STARTBDY and CONTBDY options are used. In some environments this can lead to (E)CSA fragmentation.

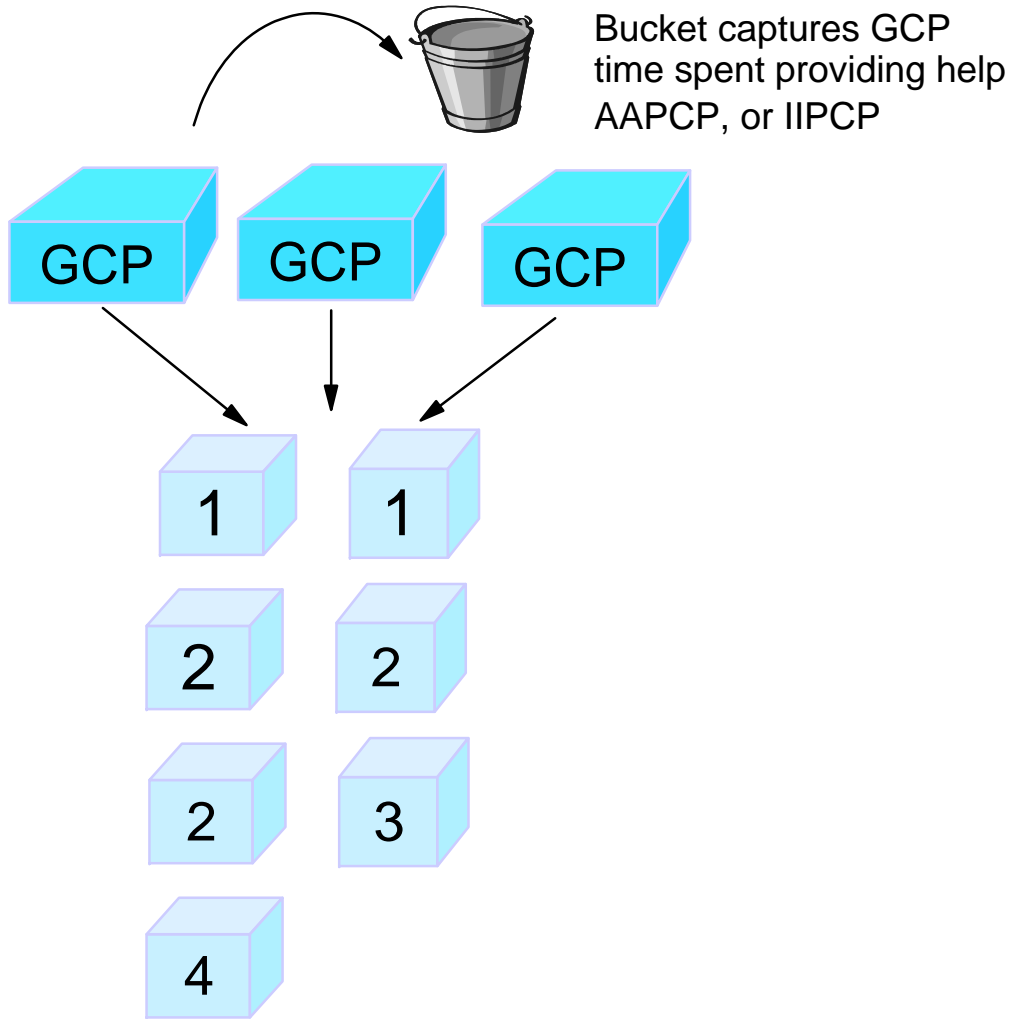
YES: Always use a "best fit" algorithm. IBM recommends you specify YES for this option to minimize (E)CSA fragmentation and to prevent user and system outages due to failed requests for (E)CSA storage.

Catalog Increase in Service Tasks

- **OA28192** - Max Number of Catalog Service Tasks
 - ▶ Number of service tasks in catalog address space can be 200 to 999
 - ▶ Hardcoded limit of 200 if maximum service tasks (999) is set
 - Override the default of 200 in SYSCATxx
 - ▶ Catalog now allows up to 400 for maximum number of service tasks
 - ▶ Each Catalog service task requires 5K of common storage (CSA/ECSA)
 - Increase of 200 to 400 will require 1M (200 * 5K) of CSA/ECSA
 - ▶ Use MODIFY CATALOG,REPORT to determine high water mark

```
*CAS*****  
* CATALOG COMPONENT LEVEL = HDZ1B10  
* CATALOG ADDRESS SPACE ASN = 0032  
* SERVICE TASK UPPER LIMIT = 180  
* SERVICE TASK LOWER LIMIT = 60  
* HIGHEST # SERVICE TASKS = 9  
* # ATTACHED SERVICE TASKS = 9
```

Management of Work - xxxHonorpriority=YES

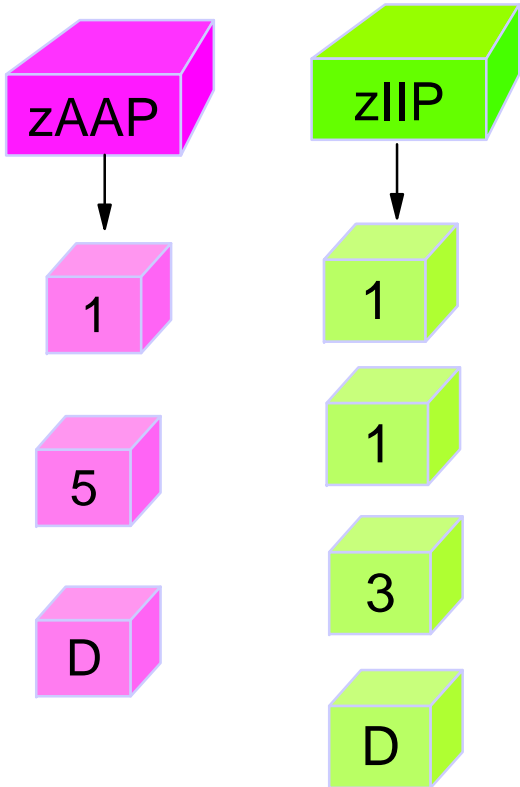


**All zAAPs Busy?
zAAP work waiting?**

YES: Signal GCP for help

**All zIIPs Busy?
zIIP work waiting?**

YES: Signal GCP for help



GCPs will help in priority order,
but won't help zIIP or zAAP work with a discretionary goal

FICON Dynamic CHPID Management

- Dynamic Channel Path Management
 - ▶ Allows system to dynamically manage channel paths in response to changing workloads
 - ▶ Balances channel and host adapter utilization
 - ▶ Simplifies I/O configuration planning and definition
 - ▶ More efficient use of hardware resources
 - ▶ Reduces channel requirements
 - ▶ Enhances Availability

- Software Dependencies - [OA28321](#)
 - ▶ All systems in LPAR cluster must be z/OS 1.11

- Hardware Dependencies
 - ▶ Processor – All currently supported processors (z900 and up)
 - ▶ Channels – All currently supported FICON channels
 - ▶ Coupling Facility required if running multi-system Switches
 - ▶ Must have control unit port (CUP) function
 - ▶ Must be defined in the IODF
 - ▶ Control Unit – No special microcode needed

- Existing ESCON DCM function is not affected

Hiperdispatch and zPCR

- zPCR using an RMF report can't determine the number of actual logical processors unparked in an LPAR when running HD=YES
 - ▶ Requires the user to manually determine the number of unparked CPs
 - ▶ RMF report values would overstate the number of logicals
 - ▶ Can't determine the unparked values for other LPARs
 - ▶ [New logical CP Optimization Capability](#)
- IRD users and hiperdispatch=NO environment the logical CPs are correctly determined by zPCR
 - ▶ IRD uses the config CPU command and actual environment is changed
 - ▶ The SMF70 record from any LPAR on a CEC can determine the actual logical CPs defined for all LPARs
- Warning message is issued when Hiperdispatch=YES systems are read into zPCR

```

-CPU 2097      MODEL 718      H/W MODEL E40      SEQUENCE CODE 00000000000999999      HIPERDISPATCH=YES
0---CPU---      ----- TIME % -----      LOG PROC      --I/O INTERRUPTS--
NUM  TYPE      ONLINE  LPAR BUSY      MVS BUSY      PARKED      SHARE %      RATE      % VIA TPI
 0   CP      100.00  91.56      91.77      0.00      100.0      3209      5.99
 1   CP      100.00  91.28      91.77      0.00      100.0      56.60     44.80
 2   CP      100.00  0.00      -----      100.00      0.0      0.00      0.00
 3   CP      100.00  0.00      -----      100.00      0.0      0.00      0.00
 4   CP      100.00  0.00      -----      100.00      0.0      0.00      0.00
 5   CP      100.00  0.00      -----      100.00      0.0      0.00      0.00
TOTAL/AVERAGE      30.47      91.77      200.0      3265      6.66

```

z/OS 1.10 Performance Items

- Contention Management Phase 3
 - ▶ Promote units of work identified by exploiters for longer periods of time
 - ▶ Promote to priority of highest-priority unit of work waiting for the held resource
 - ▶ Info on promoted workloads recorded in RMF, SDSF DA and ENC panels
- CPU Management of zIIPs
 - ▶ WLM algorithms for adjusting dispatch priorities extended to zIIP workloads
- Manage selected components in service class SYSTEM
 - ▶ Prevent inadvertent misclassification of system address spaces regardless of the WLM Policy and protect system at high utilization
 - XCFAS, GRS, SMSPDSE, SMSPDSE1, CONSOLE, IEF SCHAS, IXGLOGR, SMF, and CATALOG
- Support for 10 additional PB delay types and naming by subsystem
- Extract the WLM service definition in XML format
 - ▶ Install and activate a WLM service definition in XML format via a CIM Server
- Support for Spin and Suspend lock contention in RMF Monitor 3

HiperDispatch Mode

- PR/SM
 - ▶ Supplies topology information/updates to the z/OS guest
 - ▶ Ties high priority logicals to physicals (gives 100% share)
 - ▶ Distributes remaining share to medium priority logicals
 - ▶ Distributes any additional service to unparked low priority logicals

- z/OS
 - ▶ Associates tasks with a small subsets of logical processors
 - ▶ Dispatches work to associated subset of logicals when possible
 - ▶ Dispatches work to some other CPU when necessary
 - ▶ Parks low priority processors that are not needed or will not get service

- The combination provides the processor affinity that maximized the efficiency of the hardware caches

z/OS HiperDispatch Mode

- Improved dispatching can help reduce the effects of memory latency to improve performance and reduce CPU time
 - ▶ Better access to hardware cache
 - ▶ Manage work across fewer logical CPs
 - Correct dispatch priorities are critical
- Processors will not have equal logical share, but rather a vertical share:
 - ▶ High: 100% share
 - ▶ Medium: 0-100% share, the remainder after Vertical High share
 - ▶ Low: 0-100% share, not needed for weight and is discretionary
- IRD Vary CPU management is disabled and replaced
 - ▶ Logical CPs provided based on weight, workload demand, and available capacity
- Recommendation is to enable with the installation of the z10
 - ▶ If HD=NO ensure sufficient capacity to handle the loss of performance
- Not supported for z/OS guests under z/VM
- White Paper: <http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP101229>

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