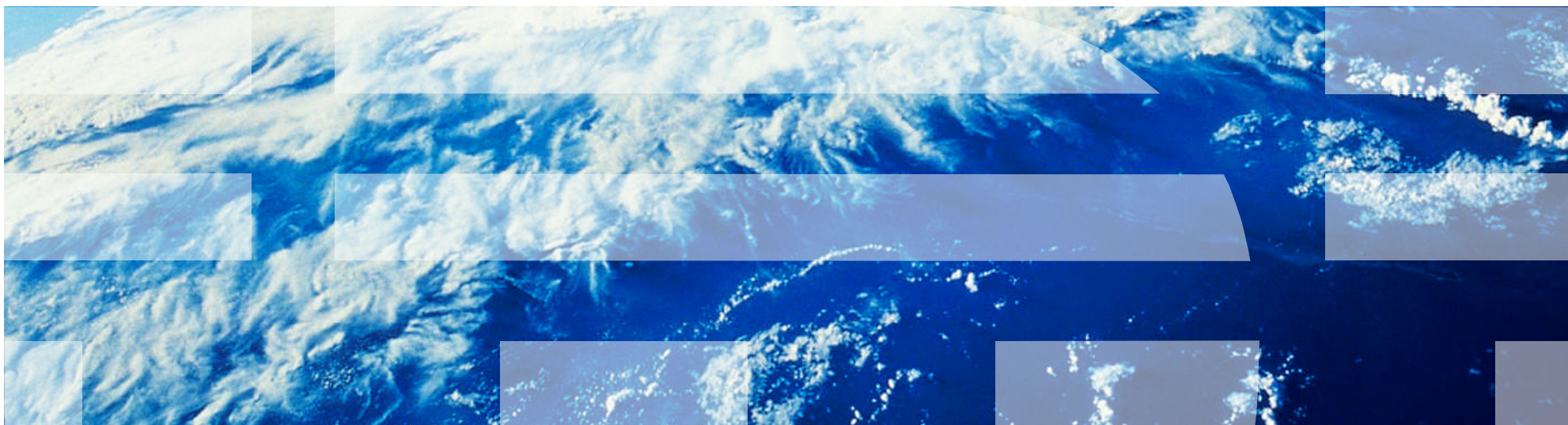


z/OS Performance **HOT** Topics

Session: 12908



Kathy Walsh
IBM Corporation



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:

*BladeCenter®, DB2®, e business (logo)®, DataPower®, ESCON, eServer, FICON, IBM®, IBM (logo)®, MVS, OS/390®, POWER6®, POWER6+, POWER7®, Power Architecture®, PowerVM®, S/390®, Sysplex Timer®, System p®, System p5, System x®, System z®, System z9®, System z10®, Tivoli®, WebSphere®, X-Architecture®, zEnterprise®, z9®, z10, z/Architecture®, z/OS®, z/VM®, z/VSE®, zSeries®

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 registered 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.

Notice Regarding Specialty Engines (e.g., zIIPs, zAAPs and IFLs):

Any information contained in this document regarding Specialty Engines ("SEs") and SE eligible workloads provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs). IBM authorizes customers to use IBM SEs only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at:

www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT").

No other workload processing is authorized for execution on an SE.

IBM offers SEs at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

Agenda

- Processor Information
 - New zEC12 Processor
 - LPAR Interrupt Delay Time
 - Warning Track
 - Relative Nest Intensity
 - CPU Measurement Facility
 - zPCR Latest Status Information
- Performance and Capacity Planning Topics
 - zBlade Capacity Planning Tool
 - WLM
 - HiperDispatch
 - DB2 and zIIPs
 - SMF
 - Other
- ▶ Addendum
 - Older APARs or Performance Information

www.ibm.com/support/techdocs

United States [change]

IBM® Search

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

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

Techdocs - the Technical Sales Library

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

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

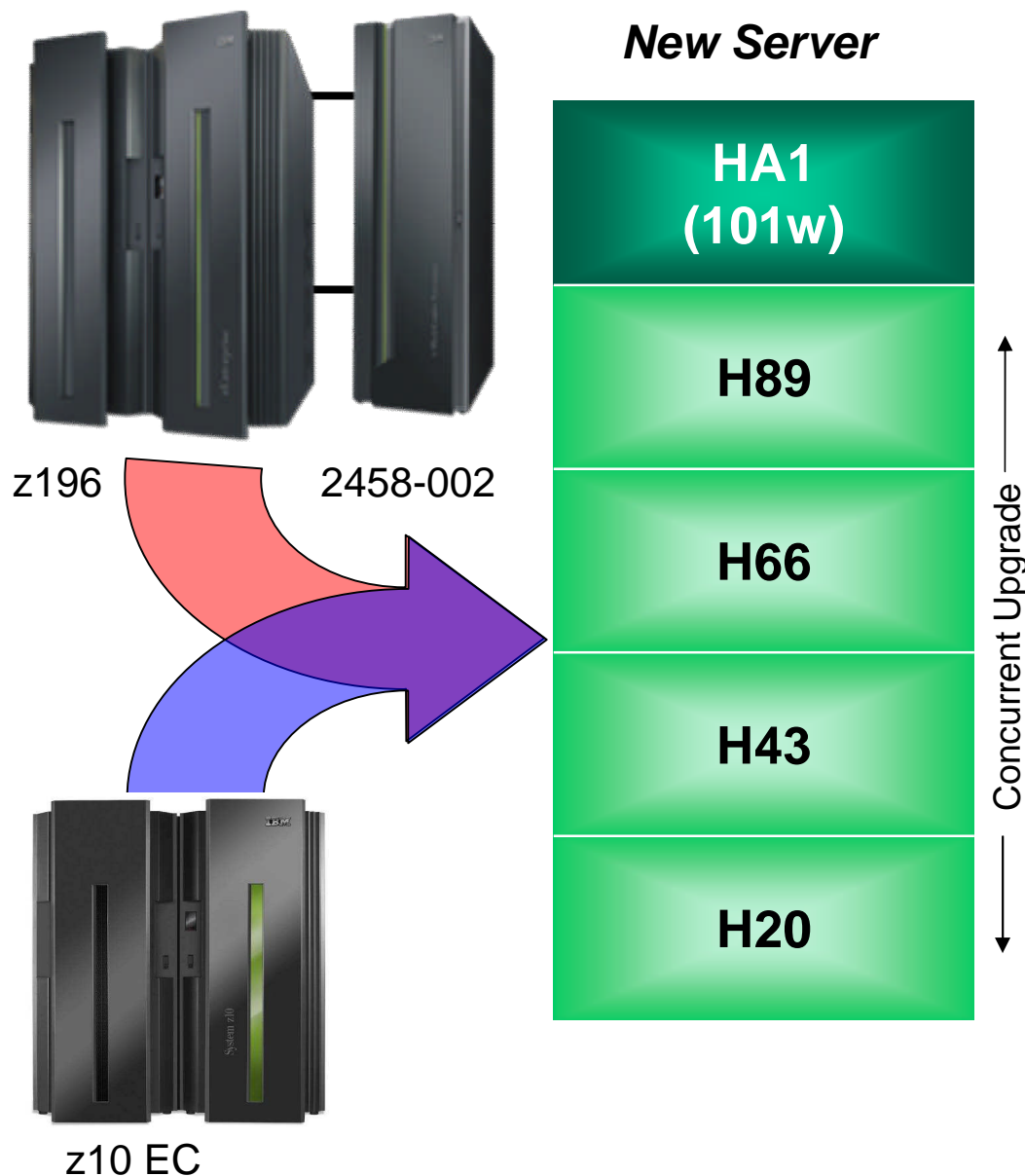
Include docs updated: [Help for Search](#)

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](#)

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

System Offering Overview



Machine Type for zEC12

- ▶ 2827

Processors

- ▶ 27 / 30 PUs per book
- ▶ Sub-capacity available up to 20 CPs
- ▶ 2 spares designated per system

Memory

- ▶ System minimum = 32 GB with separate 32 GB HSA
- ▶ Maximum 3TB / 768GB per book
- ▶ RAIM memory design
- ▶ Purchase Increments – 32, 64, 96, 112, 128, 240, 256, 512 GB

I/O

- ▶ Up to 16 connections per book (Up to 8 fanouts, 2 per fanout)
- ▶ PCIe connections 8 GB/sec
- ▶ InfiniBand 6 GB/sec

Machine Type and Model for zBX

- ▶ 2458-003

zEC12 versus z196 hardware comparison

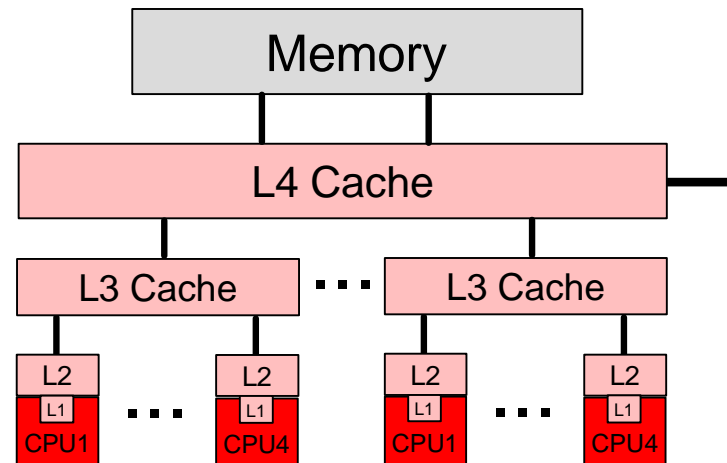
■ z196

▶ 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



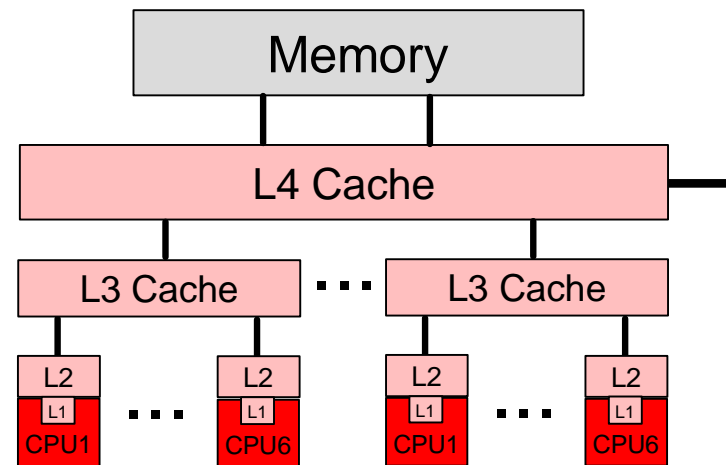
■ zEC12

▶ CPU

- 5.5 GHz
- Enhanced Out-Of-Order

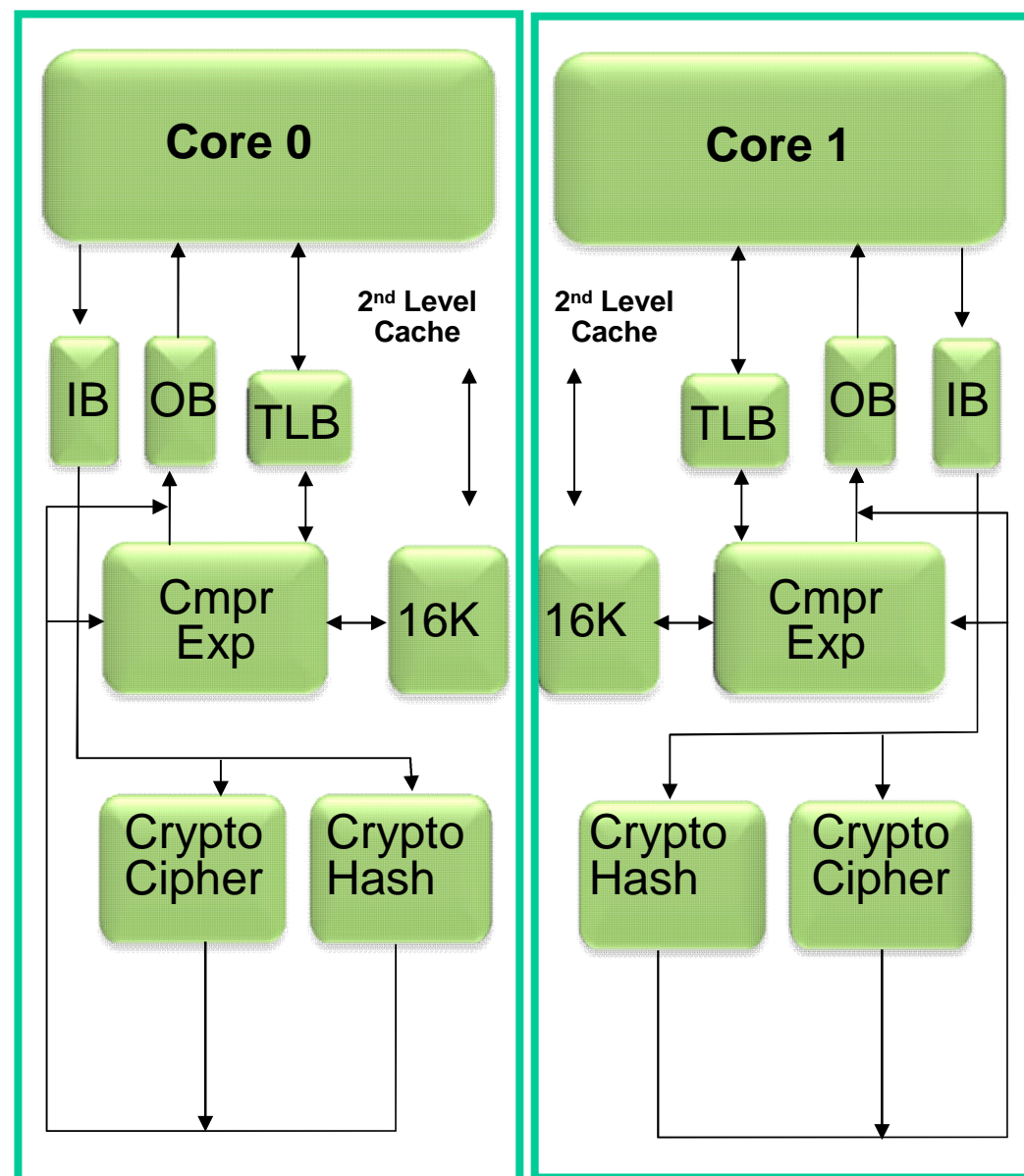
▶ Caches

- L1 private 64k i, 96k d
- L2 private 1 MB i + 1 MB d
- L3 shared 48 MB / chip
- L4 shared 384 MB / book



zEC12 Compression and Cryptographic Coprocessor

- **Coprocessor dedicated to each core (Was shared by two cores on z196)**
 - Independent compression engine
 - Independent cryptographic engine
 - Available to any processor type
 - Owning processor is busy when its coprocessor is busy
- **Data compression/expansion engine**
 - Static dictionary compression and expansion
- **CP Assist for Cryptographic Function**
 - 290-960 MB/sec bulk encryption rate
 - DES (DEA, TDEA2, TDEA3)
 - SHA-1 (160 bit)
 - SHA-2 (244, 256, 384, 512 bit)
 - AES (128, 192, 256 bit)
 - CPACF FC #3863 (No Charge) is required to enable some functions and is also required to support Crypto Express4S or Crypto Express3 features



zEC12 Architecture Extensions

- **Transactional Execution (a/k/a Transactional Memory)**
 - Software-defined sequence treated by hardware as atomic “transaction”
 - Enables significantly more efficient software
 - Highly-parallelized applications
 - Speculative code generation
 - Lock elision
- **2 GB page frames**
 - Increased efficiency for DB2 buffer pools, Java heap, others
- **Software directives to improve hardware performance**
 - Data usage intent improves cache management
 - Branch pre-load improves branch prediction effectiveness
 - Block prefetch moves data closer to processor earlier, reducing access latency
- **New Decimal-Floating-Point Zoned-Conversion Facility**
 - Improves performance of applications compiled with the new Enterprise PL/I compiler



12915: z/OS V1.13 - A Performance Update, Tuesday, 9:30 AM-10:30 AM

Transactional Execution Facility – OA38829

- Support for z/OS 1.13 running on a zEC12
- New Facility intended to be used by JAVA 7, installation code, ISV code
- Provides LOADxx external to disable use of transactional execution facility

MACHMIG – Identifies one or more facilities z/OS should not use

Column	Contents
1-7	MACHMIG
10-72	List of facilities NOT to use EDAT2 TX

- Example in LOADxx
MACHMIG TX,EDAT2

IBM Flash Express – Smarter Availability for Smarter Systems



- **Flash Express is an innovative solution designed to help you compete effectively in today's marketplace**
 - Automatically improve availability for key workloads at critical processing times
 - Drive availability and improve performance consistency for workloads that cannot afford paging spikes during key transition periods.
- Slash latency for critical application processing such as diagnostics collection
- **Extends IBM's expertise in memory management introducing a new tier of memory using Flash Express**
- **Provides a secured, resilient and immediately usable solution**
- **Planned Flash Express and pageable large page exploiters by 12/14:**
 - ***z/OS V1.13 Language Environment***
 - ***IMS 12 Common Queue Server***
 - ***Java SDK601 SR4, and Java SDK7 SR3*** and by extension exploiters such as
 - CICS Transaction Server 5.1
 - WAS Liberty Profile v8.5
 - IMS 12
 - DB2



13086: zFlash Introduction, Uses and Benefits, Tuesday, 1:30-2:30 PM

13057: zFlash Setup, Management and Configuration, Tuesday, 3:00-4:00 PM

Flash Express – What is it?

FLASH Express

- Flash Express is a PCIe IO adapter with NAND Flash SSDs
- Physically comprised of internal storage on Flash SSDs
- Used to deliver a new tier of memory- storage class memory
- Uses **PCIe I/O drawer**
- Sized to accommodate *all LPAR paging*
 - Each **card pair** provides **1.4 TB** usable storage (2.8 TB total)
 - Maximum 4 card pairs (4 X1.4=5.6 TB)
- Supported on z/OS V1.13 plus web deliverable

Designed for continuous availability

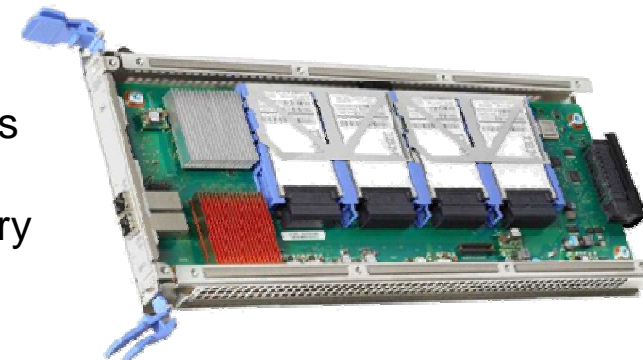
- Concurrent Firmware update for service
- RAID 10 design

Immediately usable

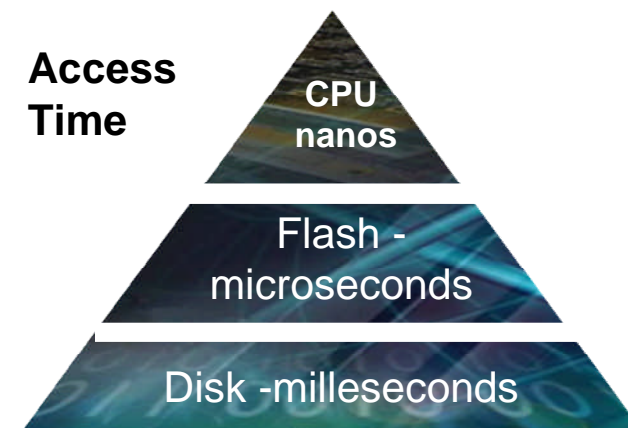
- No capacity planning needed
- No intelligent data placement needed

Secured

- Flash Express adapter is protected with 128-bit AES encryption.
- Key Management provided based on a Smart Card
- Secure Cryptographic Erase meets audit requirements



One Flash Express Card



LPAR I/O Interrupt Delay Processing

- APAR OA37160 – New Function
 - z/OS 1.12 and above
 - Requires zEC12 processor
- New Interrupt Delay Time
 - Keeps track of the time from when a subchannel is made status pending with primary status to when the status is cleared by TSCH
 - Tracking the accumulated delay encountered due to PR/SM needing to dispatch z/OS processing the interrupt as well as any z/OS delay
- APAR OA39993 - RMF support
 - RMF 74.1 record (device) and RMF 79.9
 - Enhanced RMF Post-processor device report

STORAGE	DEV	DEVICE	NUMBER	VOLUME	PAV	LCU	DEVICE	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG	AVG
GROUP	NUM	TYPE	OF	CYL	SERIAL		ACTIVITY	RESP	IOSQ	CMR	DB	INT	PEND	DISC	CONN	%	%	%	AVG	%	
							RATE	TIME	TIME	DLY	DLY	DLY	TIME	TIME	TIME	CONN	UTIL	RESV	ALLOC	ALLOC	
XTEST	2208	33903	3339	TRXSX9	1	0032	0.001	.384	.000	.128	.000	.123	.256	.000	.128	0.00	0.00	0.0	0.0	100.0	
XTEST	2209	33903	3339	TRXSXA	1	0032	0.001	.256	.000	.000	.000	.135	.256	.000	.000	0.00	0.00	0.0	0.0	100.0	
	220A	33909	10017	TRXT01	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	
	220B	33909	10017	TRXT02	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	
	220C	33909	10017	TRXT03	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	
	220D	33909	10017	TRXT04	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000	0.00	0.00	0.0	0.0	100.0	

New Function - Warning Track

OA37186 and OA37803

- **Requires zEC12**
- **PR/SM recognizes a logical CP has to be undispached from a physical CP and issues a Warning Track Interrupt (WTI, aka EXT 1007) and sets a grace period for z/OS to return the logical CP to PR/SM**
 - **If grace period expires before z/OS returns logical CP to PR/SM, PR/SM undispaches the logical CP and will redispach it later**
- **Once z/OS receives a WTI**
 - **Saves status for the work running (makes work available to dispatch on another CP)**
 - **Issues a DIAG 49C to return the CP to PR/SM**
 - **Becomes PR/SM's responsibility to redispach the logical CP when able and resume execution at instruction after DIAG 49C**
 - **z/OS keeps track of the following statistics:**
 - **How many DIAG 49Cs were successful / unsuccessful in returning CP to PR/SM before the end of the grace period**
 - **How long PR/SM undispached the logical CP for successful DIAG 49Cs**

zAAP on zIIP Support Enhancements

OA38829

- Support for z/OS 1.12 and later
- New support to allow a zAAP to be on the processor, but not in the LPAR, to allow testing of zAAP on zIIP support
- **DISPLAY IPLINFO,ZAAPZIIP,STATE** updated
 - **IEE256I ZAAPZIIP STATE: *state***
 - In the message text, *state* is one of the following
 - ACTIVE: zAAP on zIIP is active
 - INACTIVE – ZAAPZIIP SYSTEM PARAMETER IS 'NO'
 - INACTIVE – NO ZIIP(S) DEFINED TO THIS LPAR
 - INACTIVE – ZAAP(S) DEFINED TO THIS LPAR
 - INACTIVE – ZAAP(S) INSTALLED ON THE MACHINE ← ***deleted !***
 - INACTIVE – TOO MANY ZAAPS+ZIIPS INSTALLED ON THE MACHINE ← ***new !***
 - INACTIVE – GLOBAL MACHINE DATA IS NOT AVAILABLE TO THIS LPAR
- Statement of Direction: IBM zEnterprise EC12 is planned to be the last high-end System z server to offer support for zAAP specialty engine processors.

CPU Measurement Facility

- Hardware Instrumentation Facility available on z10 GA2, z196, z114, and zEC12
 - n New z/OS component - Hardware Instrumentation Facility (HIS)
 - n Generates SMF 113.2 records
 - n z/VM support for CPU MF Counters via **APAR VM64961**
 - z/VM 6.1 and z/VM 5.4 on z10s and later servers

- Capacity Planning Changes
 - n Data is now deemed as critical when doing System z capacity planning
 - n When available zPCR will use the SMF 113 data to select workload

- RNI calculations for z196 and z114 changed slightly (6/2012)
 - n Refinement based on customer measurements and LSPR workloads
 - n No change to z10 RNI calculation

- RNI calculations for zEC12 published on Techdocs (9/2012)
 - n [Http://www.ibm.com/support/techdocs/atmastr.nsf/WebIndex/TC000066](http://www.ibm.com/support/techdocs/atmastr.nsf/WebIndex/TC000066)

13094: The Relatively New LSPR and zEC12, Wednesday, 9:30 AM-10:30 AM

13098: CPU MF - 2013 Update and WSC Experiences, Wednesday, 1:30 PM-2:30 PM

New CPU Measurement Education – PRS4922

- WSC hosted two Part Webinar to provide CPUMF Education
 - Part 1
 - Introduction and Overview of CPU MF
 - Implementation Details
 - Solicited customer data to be used in Part 2
 - Each customer sent in SMF 113 records and received a customized report
 - Part 2
 - Approx. 40 customers responded with data
 - Detailed information on usage and meaning of the CPUMF metrics
 - Overview and profile information on the provided data

- Presentations, and replays are available
 - <http://www-03.ibm.com/support/techdocs/atstr.nsf/WebIndex/PRS4922>

Looking for zEC12 Migration “Volunteers” to send SMF data

- Want to validate / refine Workload selection metrics

Looking for “Volunteers”

(3 days, 24 hours/day, SMF 70s, 72s, 113s per LPAR)

“Before z10 / z196” and “After zEC12”

Production partitions preferred

If interested send note to jpburg@us.ibm.com,

No deliverable will be returned

Benefit: Opportunity to ensure your data is used to influence analysis

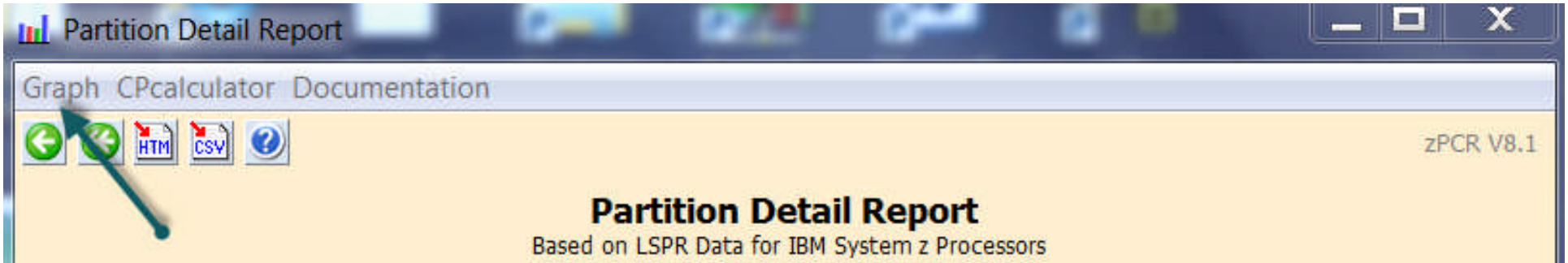
zPCR Latest Status

- Version 8.1 (9/2012)
 - IBM zEnterprise EC12 (zEC12) has been added
 - LSPR data is now based on z/OS 1.13
 - zAware can be configured as an SCP type in either the GCP or IFL pool

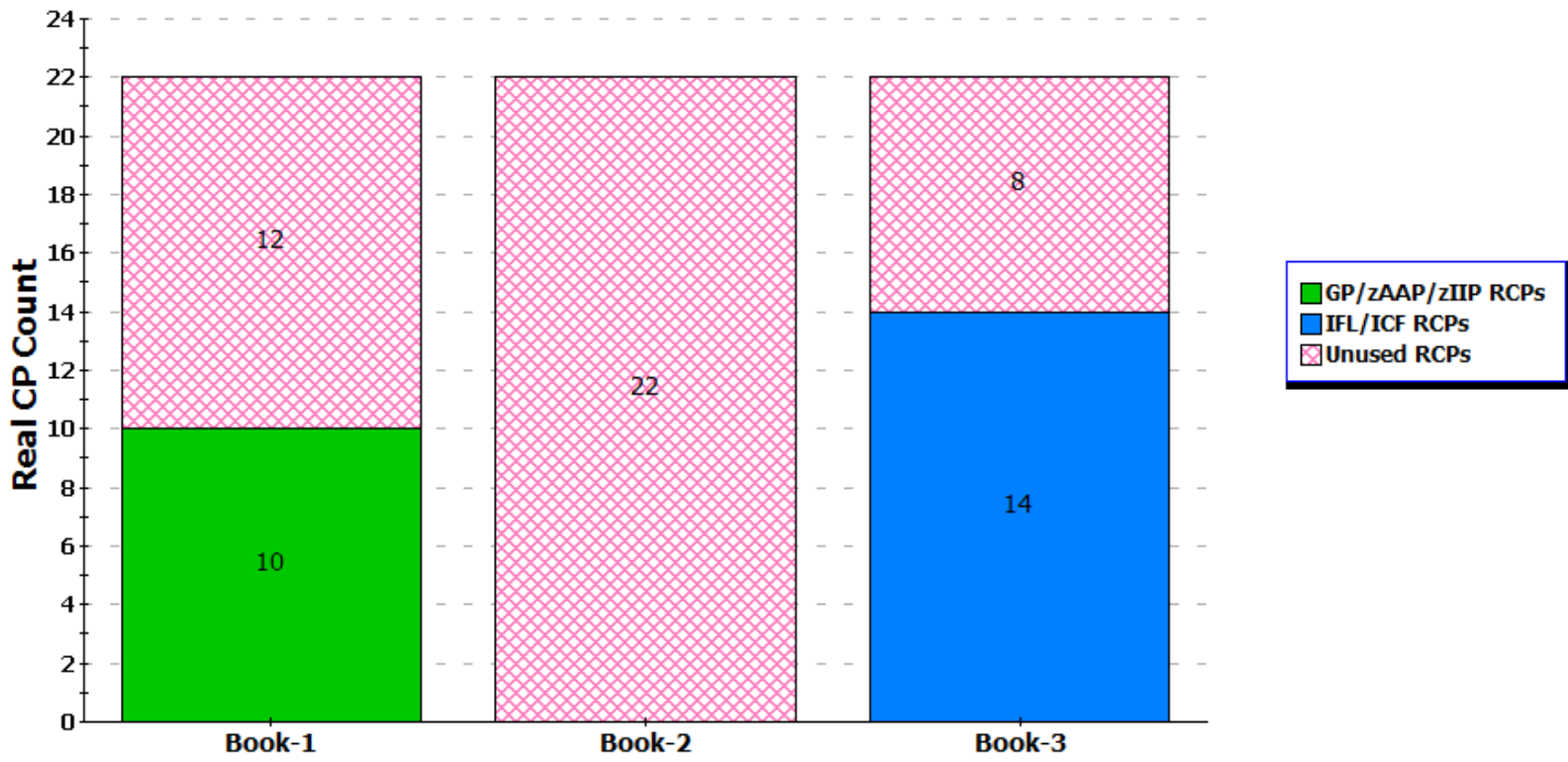
- Version 7.9b (7/31/12) New Functions
 - **Algorithms concerning System z Multi-Book models have been enhanced**
 - Capacity results will likely differ from previous zPCR versions
 - zPCR algorithms for multi-book configurations are updated to recognize the order in which real CPs get allocated across books
 - GP, zAAP, and zIIP CPs are allocated upward, starting in the first book, IFL and ICF CPs are allocated downward, starting in the last book
 - An intersection of GP/zAAP/zIIP CPs with IFL/ICF CPs can occur in only one book and zPCR applies the usual partitioning cost for CPs in that book only, while cost for CPs in remaining books is less

13097: zPCR Capacity Sizing Lab - Part 1: Intro and Overview, Thursday, 1:30 PM-2:30 PM

12674: zPCR Capacity Sizing Lab - Part 2: Hands-on Lab, Thursday, 3:00 PM-4:00 PM

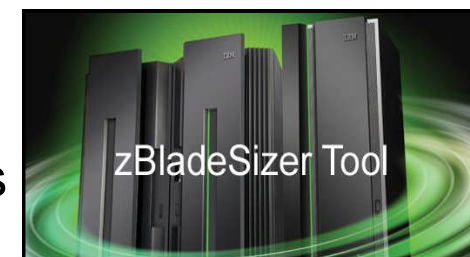


Real CP Distribution Across Books
LPAR Host: 2827-H66/700
10-GP/zAAP/zIIP; 14-IFL/ICF



No RCPs in contention

zBX Capacity Planning Sizing Tool



- IBM zBladeEXTR
 - Identify servers and time periods to consider in the analysis
 - Reads distributed servers performance data as input
 - Run by the installation to reduce amount of data needed to be sent to IBM
- IBM zBladeSizer
 - Provides optimal solution into a zBX Mod 002 including CPU, memory, network I/O, and disk I/O based on configuration constraints
 - Allows definition of availability and performance criteria for the environment
 - Supports
 - **POWER7™** blades
 - **System x®** blades - Linux servers only
- Supports new and existing zBX environments
- Contact your IBM Account team to arrange a sizing

13091: zBX Capacity Sizing using IBM zBladeSizer and IBM zBladeEXTR, Thurs, 4:30-5:30

12946: Application Performance Management and Capacity Planning for IBM zEnterprise Hybrid Workloads, Monday, 4:30 PM-5:30 PM

WLM – OA32298 – New Function

- New I/O Priority Manager feature in the IBM System Storage DS8700 and DS8800 provides favored processing of selected I/O requests by throttling other I/O requests which are less important
- Controlled by new IEAOPTxx parameter
 - STORAGESEVERMGT = YES | NO
 - Requires IO Priority Management in WLM Policy to be set to YES
 - Specifying YES removes control unit delay samples from service classes with velocity goals
 - Achieved velocity may change if significant control unit queue delay (part of PEND time)
- Activated by WLM passing an I/O mgmt field with I/O request to storage server
 - WLM provides following information by service class period:
 - Response Time Goal: Goal achievement (derived from PI) and specified importance
 - Velocity Goal: Specified velocity goal and importance
 - System Goals: Not managed by Server I/O Priority Manager
 - Discretionary Goal: Always eligible to be throttled by Server I/O Priority Manager
- White Paper: IBM System z DS8000 I/O Priority Manager
<http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP102074>

HiperDispatch APARs

- OA37736 – New Function
 - WLM uses free processor capacity of the CEC to determine if an LPAR should unpark one or more vertical lows
 - Apportionment didn't take into account relative weight of LPARs on CEC
 - Small LPARs could unpark and take free processor capacity leaving a larger LPAR unable to unpark vertical lows
 - LPARs may then not have proportional access, based on weight, to whitespace
 - Changes:
 - Now unpark an additional vertical low if CPU consumption of the partition is below its current weight **plus** newly introduced unused capacity share
 - Unparking only happens when no free processor capacity and the LPAR has CPU demand
 - SMF 99.12 and 99.14 record mappings will be published
 - 99.12 – Hiperdispatch Interval data
 - 99.14 – Hiperdispatch Topology data

Workload Promotion – LCK – Local Lcok

- In HiperDispatch while a WEB is suspended while holding a local/CML lock, z/OS will promote the WEB to dispatch priority x'FF'
- Done so work will run at a high priority until it releases its Local/CML lock
- Monitors will not display the WLM designated Dispatch Priority
- May introduce CPU delay to high importance workloads

WORKLOAD ACTIVITY

--PROMOTED--

BLK	0.000
ENQ	0.000
CRM	0.000
LCK	0.275
SUP	0.000

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

z/OS 1.10 RMF Enhancements

- RMF Monitor III provides detailed information about spin and suspend locks
- Suspend Lock report will display the address spaces which hold locks and which are suspended
- Spin Lock report will display how often global locks are held and who is spinning
- Specify LOCK | NOLOCK in RMF Monitor III parmlib
- Reporting of lock statistics is intended to help analyze lock contention in the system

z/OS 1.13 RMF GRS & Supervisor Delay Monitoring

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

12792: Remote RMF Report Access - Hands-on Lab, Monday, 3:00 PM-4:00 PM

13089: RMF: The Latest and Greatest, Thursday, 8:00 AM-9:00 AM

Enhanced SMF 30 Reporting

- APAR OA39629 – New Function
- New SMF 30 function to provide enhanced reporting in the CPU Accounting section
 - z/OS 1.12 and above
- Highest percent of CPU time used by a single task in the address space in the interval, step, or job
 - SMF30_Highest_Task_CPU_Percent
- Program name associated with the task with the highest percentage of CPU time in the address space
 - SMF30_Highest_Task_CPU_Program

New DB2 / WLM Goal Setting Documentation

- Updated WLM classification information for DB2 started tasks
 - New section published June, 2012

- Recommends:
 - IRLMPROC in SYSSTC
 - *ssnm*MSTR, *ssnm*DBM1, *ssnm*DIST and WLM-managed stored **procedure address spaces** should be in a service class with a high importance and an aggressive velocity goal
 - May need to be marked CPU Critical
 - Recommends LPARs which run DB2 work at low dispatch priorities in environment with high CPU utilization to be more aggressive with blocked workload support
 - Protects environment by ensuring work holding DB2 locks gets CPU service
 - Recommends setting BLWLINTHD in IEAOPTxx from default of 20 to 3-5 secs

Predicting DB2 10 zIIP Eligibility for Prefetch and Deferred Writes

- In DB2 V9 workloads executing under 'non-preemptable' SRBs in DBM1 include:
 - Prefetch
 - Deferred write
 - Castout
 - Pseudo close
 - Write scheduling
- In DB2 10, prefetch and deferred write are executed under 'preemptable' SRB processing in DBM1
 - These workloads are marked as zIIP eligible
 - Represents the vast majority of the DB2 9 'non-preemptable' SRB processing in DBM1 address space
 - Will depend on workload mix

- Be sure you have sufficient zIIP capacity and DB2 address spaces are appropriately defined to
 - Time sensitive activities
- If you have 5 GCP and 1 zIIP you have 1/5th the opportunity to get dispatched
- May need to plan increased zIIP capacity when moving to DB2 10

V9 and V10 Sample CPU Times from SMF 100

■ V9 example:

CPU,TIMES	TCB TIME	PREEMPT SRB	NONPREEMPT SRB	TOTAL TIME	PREEMPT IIP SRB	/COMMIT
SYSTEM SERVICES ADDRESS SPACE	6.186504	0.001054	13.465251	19.652809	N/A	0.001272
DATABASE SERVICES ADDRESS SPACE(DBM1)	0.292939	0.013103	41.596518	41.902560	0.000000	0.002712
IRLM	0.000128	0.000000	0.420866	0.420994	N/A	0.000027
DDF ADDRESS SPACE	0.003241	0.000000	0.005809	0.009051	0.000000	0.000001

■ V10 example:

CPU,TIMES	TCB TIME	PREEMPT SRB	NONPREEMPT SRB	TOTAL TIME	PREEMPT IIP SRB	/COMMIT
SYSTEM SERVICES ADDRESS SPACE	5.388198	9.265621	1.226727	15.880545	N/A	0.001083
DATABASE SERVICES ADDRESS SPACE	0.144509	0.838230	0.022309	1.005048	25.877390	0.000069
IRLM	0.000060	0.000000	0.266525	0.266585	N/A	0.000018
DDF ADDRESS SPACE	0.086311	0.000138	0.004407	0.090856	0.000000	0.000006
TOTAL	5.619079	10.103988	1.519967	17.243034	25.877390	0.001176

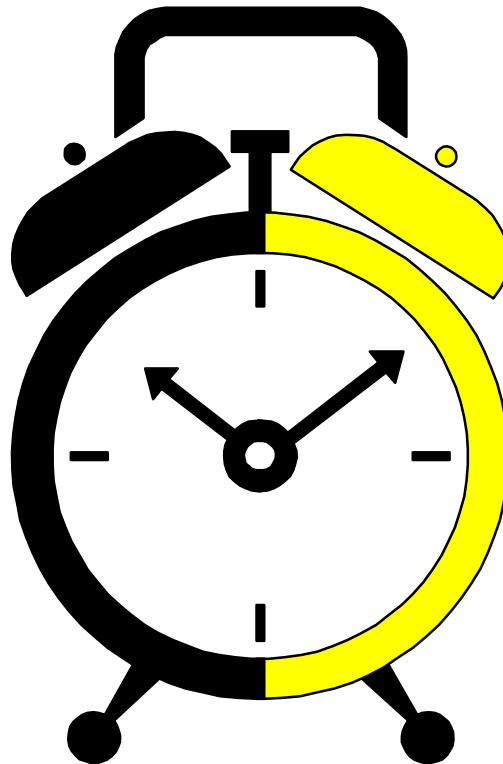
Example of Potential Offload

- Eight Data Sharing Members during Online Peak

Total DB2 Started Task Avg. GCPs used minus DDF Threads	DBM1 NonPreempt Avg. GCPs	Total DBM1 Avg. GCPs	Estimated Average zIIP CPs Offloaded in V10	Estimated zIIP Offload as % of Total DB2 Started Tasks
1.08	0.49	0.57	0.49	46%
1.56	0.66	0.80	0.66	42%
1.01	0.41	0.50	0.41	41%
0.62	0.52	0.52	0.52	82%
0.30	0.26	0.26	0.26	86%
0.10	0.08	0.08	0.08	77%
1.08	0.51	0.58	0.51	47%
0.84	0.29	0.49	0.29	34%

Addendum

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



New Enclave Based Workloads

- Exploiters are continually making additional use of both independent and dependent enclaves
 - Watch release level migrations to ensure new ones are properly classified
 - Default classification is SYSOTHER
- Use a monitor like SDSF's enclave panel to check
 - LDAP
 - System Rexx
 - TCPIP IPSEC
 - GRS Monitor

<u>SDSF ENCLAVE DISPLAY</u>							SYSD	ALL	LINE 1-12
COMMAND INPUT ==>									SC
NP	NAME	<u>SSType</u>	Status	<u>SrvClass</u>	Per	PGN	<u>RptClass</u>		
	3C000F54BE	LDAP	INACTIVE	OPS_HIGH	1		RLDAPPRB		
	5C000F54C0	LDAP	INACTIVE	OPS_HIGH	1		RLDAPDEF		
	60000F54BD	LDAP	ACTIVE	OPS_HIGH	1		RLDAPGEN		
	2000000001	STC	INACTIVE	SYSTEM	1				
	3800007944	STC	INACTIVE	SYSSTC	1				
	2800000003	STC	INACTIVE	SYSSTC	1		RTCPIP		
	2C00000004	TCP	INACTIVE	SYSOTHER	1				

WP101229 - 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 pattern
 - 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

■ OA36054

- Beginning with z/OS 1.13 when running on an IBM zEnterprise z196, z114, or zEC12 the default for Hiperdispatch will be YES

Share of the partition - assumes 1.5 logical to physical ratio	Number of Physical CPs + zIPs + 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