
z/OS Performance **HOT** Topics

Session: 10594

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:

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

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries. Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

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

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed.

Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

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

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Agenda

- Processor Information
 - IBM® zEnterprise™ 114 (z114)
 - CPU Measurement Facility
 - Power Saving Mode
 - zPCR Information
- Performance and Capacity Planning Topics
 - WLM
 - HiperDispatch
 - DB2 and zIIPs
 - Other
- New z/OS Performance Support Overview
 - z/OS 1.13
 - z/OS 1.12
- ▶ Addendum
 - Older APARs or Performance Information

www.ibm.com/support/techdocs

IBM United States [change]

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

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

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: "Fuzzy" search
 Hits: Order by:
 Include docs updated: [Help for Search](#)

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

Related links

- Redbook publications
- IBM Software Support Handbook

Techdocs Library

- Flashes
- Presentations & tools
- Technotes & tips
- FAQs
- White papers
- Solution scenario profiles
- Customer support plans
- Sizings
- Auxiliary Material
- Search Techdocs
- Techdocs feedback

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?

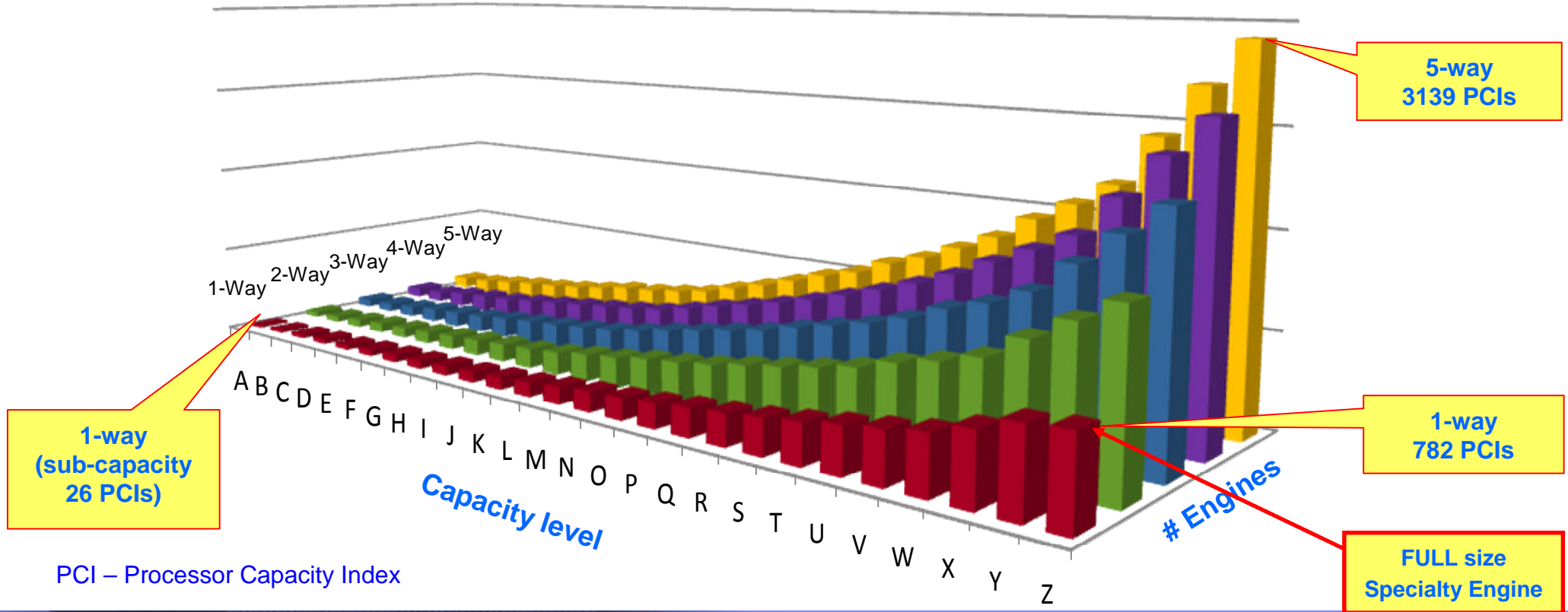
Looking for support resources or other documents and tools?

→ [Support & downloads](#)

zEnterprise 114

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

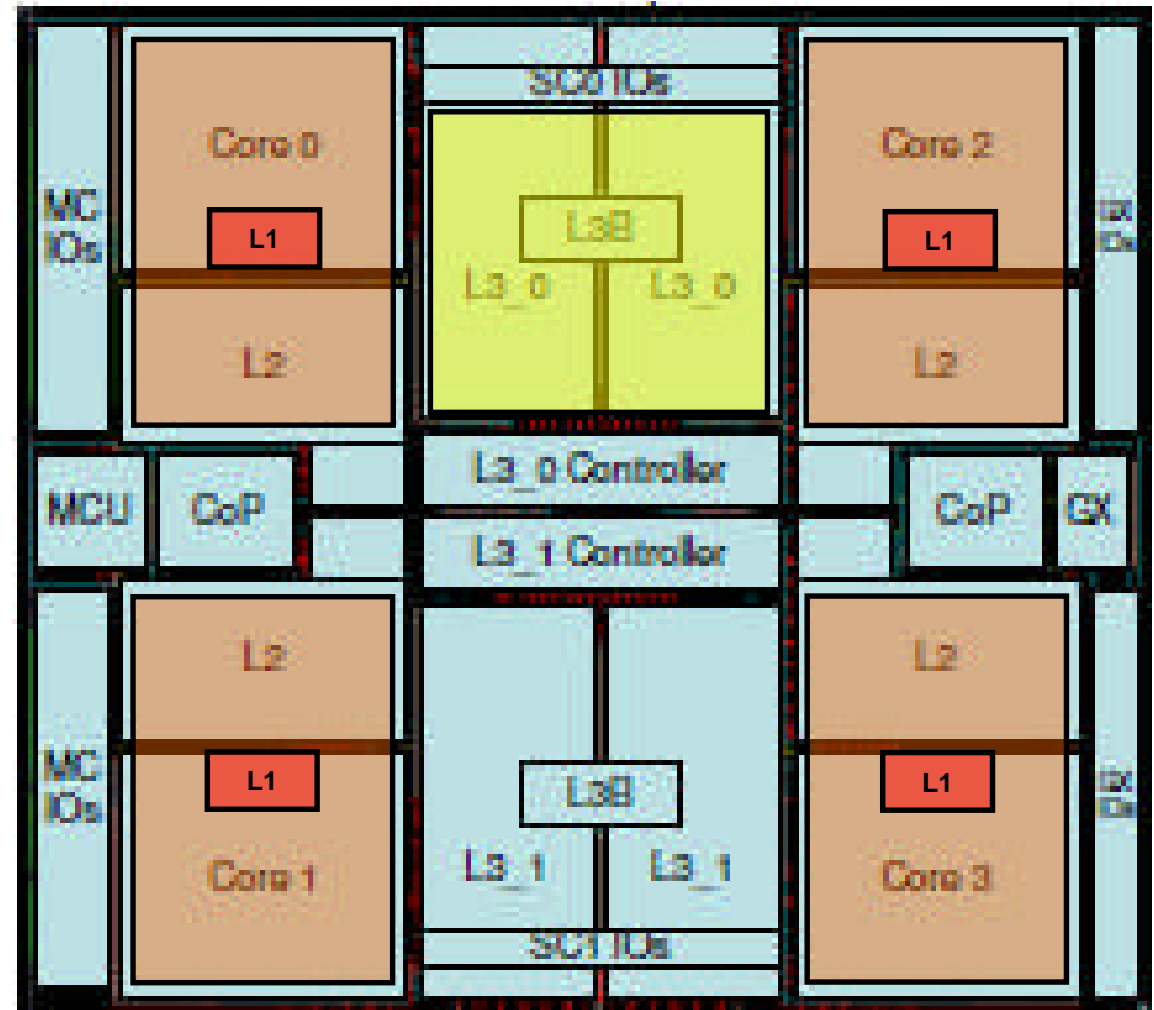
Number of z114 CPs	Base Ratio	Ratio z10 BC to z114
1 CP	z10 BC Z01	1.18
2 CPs	z10 BC Z02	1.16
3 CPs	z10 BC Z03	1.14
4 CPs	z10 BC Z04	1.13
5 CPs	z10 BC Z05	1.12



PCI – Processor Capacity Index

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

- Quad core chips with 3 or 4 active cores
 - Same as the zEnterprise 196
- 3.8 GHz
- L1: 64K I / 128K D private/core
- L2: 1.5M I+D private/core
- L3: 12MB
 - Same chip as z196, but enabled half of the available 24MB
- L4: 96MB per processing drawer
 - On the SC Chip
 - 24MB assigned to each core
 - $24 \times 4 = 96$
 - Half of that on the z196 (192MB)



Components of Workload Capacity Performance

■ Instruction Complexity (Micro Processor Design)

- Many design alternatives
 - Cycle time (GHz), instruction architecture, pipeline, superscalar, Out-Of-Order, branch prediction, and more
- Workload effect
 - May be different with each processor design
 - **But once established for a workload on a processor, doesn't change very much**

■ Memory Hierarchy or “Nest”

- Many design alternatives
 - Cache (levels, size, private, shared, latency, MESI protocol), controller, data buses
- Workload effect
 - Quite variable
 - **Sensitive to many factors: locality of reference, dispatch rate, IO rate, competition with other applications and/or LPARs, and more**
 - Net effect of these factors represented in “Relative Nest Intensity”
- **Relative Nest Intensity (RNI)**
 - **Activity beyond private-on-chip cache(s) is the most sensitive area**
 - **Reflects distribution and latency of sourcing from shared caches and memory**
 - Level 1 cache miss per 100 instructions (L1MP) also important
 - Data for calculation available from CPU MF (SMF 113) starting with z10

CPU Measurement Facility

- Hardware Instrumentation Facility available on z10 GA2, z196, and the z114
 - New z/OS component - Hardware Instrumentation Facility (HIS)
 - Generates SMF 113.2 records
 - z/VM support for CPU MF Counters via **APAR VM64961**
 - **z/VM 6.1 and z/VM 5.4 on z10s and z196s**
- New Support Information
 - **OA36816** - Automates the process to allow HIS to continue running while indicating data loss in the SMF 113 records for the interval
- Capacity Planning Changes
 - This data needs to be collected and used to select appropriate workload when doing capacity planning
 - When available zPCR will use the SMF 113 data to select workload

Session 10886: CPU MF - 2012 Update and WSC Experiences, Wed, 3:00 PM

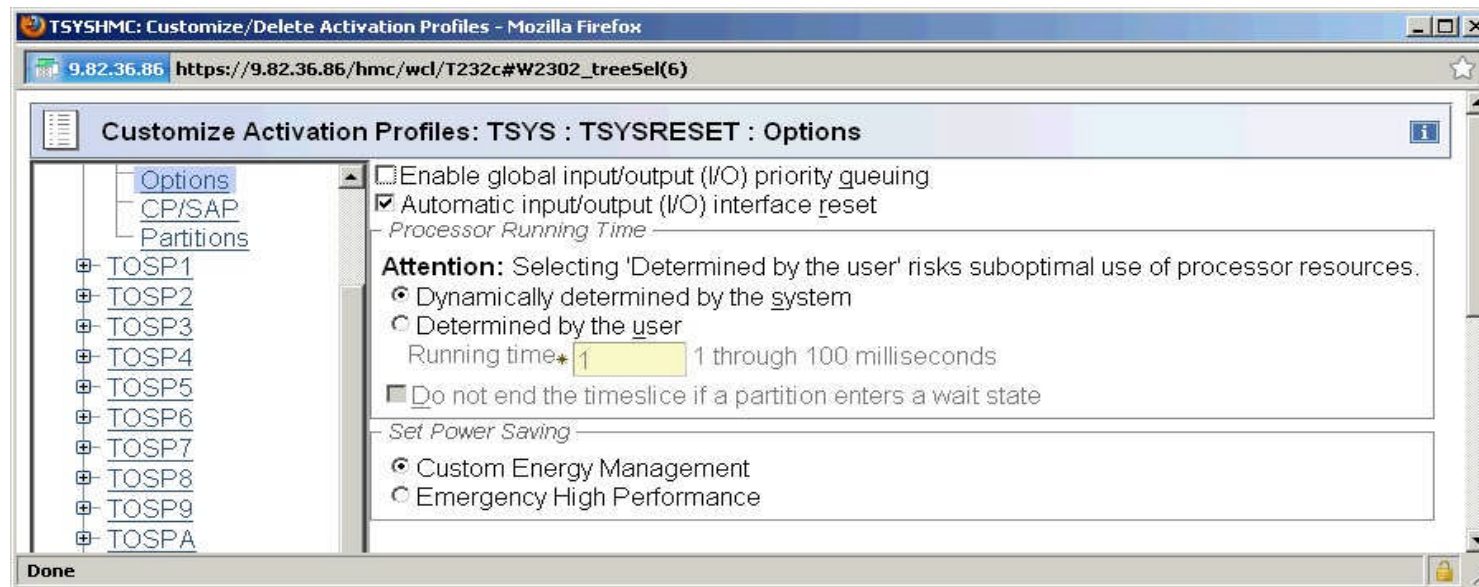
zPCR Latest Status

- Version 7.7 (1/2012) New Functions
 - Easier LPAR Configuration Input via CP3KEXTR
 - Program will post-process SMF 113 records (CPU-MF) and be used to identify appropriate LSPR workload for each LPAR
 - A Remove Parked LCPs from Partition LCP Count checkbox is now provided
 - New support for alternative LCP count settings for shared GP partitions
 - Unparked LCPs only (as read from EDF or RMF)
 - Moderate or Minimum (based on partition weights)
 - User defined overrides
 - A default workload category for that SCP will be assigned
 - If SMF 113s provided via EDF this will be used to set information
 - If DASD IO information provided via EDF then this will be used to set information
 - No supporting information provided then default workload category for that System Control Program (SCP) will be assigned
 - User's Guide and Online Help have been updated

Session 10885: zPCR Capacity Sizing Lab - Part 1 Introduction and Overview, Thur, 11AM
10880: zPCR Capacity Sizing Lab - Part 2 Hands-on Lab, Thur, 12:15 PM

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



Full Power

TSYS Details - TSYS

Instance Information	Acceptable Status	Product Information	Network Information
----------------------	-------------------	---------------------	---------------------

CPC

Power rating: 65644 W
 Power consumption: 32162 W
 Power saving: High performance
 Power capping: Custom
 Cap range: 44409 W - 106936 W
 Current cap: 106936 W

zCPC

Power rating: 27400 W
 Power consumption: 18937 W
 Ambient temperature: 16.4°C (61.5°F)
 Exhaust temperature: 30.0°C (86.0°F)
 Humidity: 34 %
 Dew point: .5°C (32.9°F)
 Heat load: 64661 BTU/hr.
 Heat load (forced-air): 64661 BTU/hr.
 Heat load (water): 0 BTU/hr.
 Maximum potential power: 21351 W
 Maximum potential heat load: 72901 BTU/hr.
 Power saving: High performance
 Power capping: Enabled
 Cap range: 21351 W - 27400 W
 Current cap: 27400 W

Power Save Mode

TSYS Details - TSYS

Instance Information	Acceptable Status	Product Information	Network Information
----------------------	-------------------	---------------------	---------------------

CPC

Power rating: 65644 W
 Power consumption: 29275 W
 Power saving: Custom
 Power capping: Custom
 Cap range: 44409 W - 106936 W
 Current cap: 106936 W

zCPC

Power rating: 27400 W
 Power consumption: 15933 W
 Ambient temperature: 15.2°C (59.4°F)
 Exhaust temperature: 27.0°C (80.6°F)
 Humidity: 36 %
 Dew point: .4°C (32.7°F)
 Heat load: 54402 BTU/hr.
 Heat load (forced-air): 54402 BTU/hr.
 Heat load (water): 0 BTU/hr.
 Maximum potential power: 21351 W
 Maximum potential heat load: 72901 BTU/hr.
 Power saving: Low power
 Power capping: Enabled
 Cap range: 21351 W - 27400 W
 Current cap: 27400 W

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  INTERVAL 03.17.221
                                RPT VERSION V1R12 RMF          TIME 00.20.00  CYCLE 1.000 SECONDS

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

```


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
 - STORAGESERVERMGT = 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
 - May see change in achieved velocity when enabled if significant control unit queue delay (part of PEND time)
- Activated by WLM passing an I/O management field with the I/O request to the storage server
 - WLM sets this field for each address space and enclave and its value will be propagated by IOS when it sends the I/O request to the storage servers
 - 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/atsmastr.nsf/WebIndex/WP102074>

CICS Response Time Management Enhancement

- OA35428 (WLM) and OA35617 (SMF) - New Function
 - WLM manages CICS regions solely based on the transaction service class mix being executed
 - Depending on mix, WLM is generally not able to distinguish CICS TORs from AORs
 - CICS TORs require short, very fast access to resources and typically are not CPU intensive
 - CICS AORs typically run more CPU intensive work which do not require the same fast access as TORs
 - At higher utilization levels, >85% busy noticeable queue (QRMod) delays within the TORs can be recognized
 - Increases end-to-end response times of the CICS transactions, and reduces the throughput of CICS work
 - Of most concern in exclusively CICS environments with little postponeable work
 - More visible in a HIPERDISPATCH=YES environment
 - Number of logical processors in use is minimized in order to increase cache hits and thereby throughput of the work in the system

CICS Response Time Management Enhancement

- Possible circumvention:
 - Exempting CICS TORs from being managed using the goals of the transaction service classes
 - CICS TORs then managed towards the velocity goals
- Drawback is transaction statistics not available for managing CICS work
- New Function:
 - New WLM Classification option *BOTH* for managing CICS TOR regions
 - TOR managed to velocity goals of the region but transaction completions are still tracked so WLM can manage CICS service classes with response time goals
 - Option *BOTH* should only be used for CICS TORs with noticeable queue delays at higher utilization levels
 - CICS TORs should be defined with a higher importance than the response time service classes for the CICS transactions
 - CICS AORs should remain at the default TRANSACTION
 - SMF 30 record updated with new bit, SMF30CRM, to indicate address space matched a classification rule of '*BOTH*'

CICS Response Time Management Enhancement

- New WLM BOTH classification support causes the function introduced in APAR OA34801 to be obsolete
 - OA34801 introduced support for IEAOPT parameter REPORTCOMPLETIONS
 - Specifies whether to allow transaction response time data to be reported for transaction service classes even if the subsystem work manager region is exempted from being managed towards the transaction goal
 - Default value is NO
 - Function is removed in z/OS V1R13 and higher
 - Function in z/OS V1R11 and z/OS V1R12 is still supported but not recommended and new WLM Classification option BOTH is the recommended solution

z/OS LDAP Tivoli Directory Services and WLM

- OA36644 – Provides ability to define performance goals for work within the LDAP server
 - Support is always active and a default service class needs to be defined in the LDAP classification subsystem
 - Failure to do this will result in LDAP work running in SYSOTHER
 - Creates a never ending independent enclave
 - Performance goals are set using:
 - Client's IP address, bound user's distinguished name, both the IP address and the distinguished name, request matching a search pattern in the operations monitor
 - Using either wlmExcept configuration options, or WLMEXCEPT operator command a client request can be mapped to a 'Transaction Name'
 - Use the WLM TN classification attribute to assign a service class
 - Default Transaction Name called 'GENERAL' is provided
 - Migration Action: LDAP subsystem name was used previously in WLM so verify the rules are as you intended

z/OS LDAP Tivoli Directory Services and WLM

- Required to provide WLM rule for TN General or a default service class
 - Don't create a special purpose service class for this work
 - Question is where to put it

```

-----
                          Modify Rules for the Subsystem Type                Row 1 to 2 of 2
Command ===> _____ Scroll ===> CSR

Subsystem Type . : LDAP                Fold qualifier names?   N   (Y or N)
Description . . . LDAP Tivoli Directory Services

Action codes:   A=After      C=Copy      M=Move      I=Insert rule
                B=Before    D=Delete row R=Repeat  IS=Insert Sub-rule
                                           More ===>

Action      -----Qualifier-----          -----Class-----
Action      Type      Name      Start      Service      Report
_____ 1  TN      IPWEBPRB  _____  OPSHI        _____
_____ 1  TN      GENERAL  _____  OPSHI        _____

          DEFAULTS: OPSHI
    
```

z/OS LDAP Tivoli Directory Services and WLM

- Required to provide WLM rule for Transaction Name 'General' or a default service class
 - Don't create a special purpose service class for this work
 - Create a report class
 - Question is where to put it in the WLM Classification schema
- An independent enclave BUT it's a never ending enclave
 - Can't use a response time goal, must use velocity
 - Can't specify SYSSTC where LDAP may have been before
 - Normally don't mix enclave and non-enclave work in a service class

```
REPORT BY: POLICY=WSCDEF      WORKLOAD=ONL_WKL      SERVICE CLASS=DNSATAACK
                                CRITICAL      =NONE
-TRANSACTIONS-  TRANS-TIME  HHH.MM.SS.TTT  --DASD  I/O--  ---SERVICE---
AVG          1.00  ACTUAL          0  SSCHRT   0.0  IOC          0
MPL          1.00  EXECUTION          0  RESP    0.0  CPU         3329
ENDED        0    QUEUED            0  CONN    0.0  MSO          0
END/S        0.00  R/S AFFIN          0  DISC    0.0  SRB          0
#SWAPS       0    INELIGIBLE         0  Q+PEND  0.0  TOT         3329
EXCTD        0    CONVERSION         0  IOSQ    0.0  /SEC         28
AVG ENC      1.00  STD DEV            0
REM ENC      0.00
MS ENC       0.00
                                ABSRPTN     28
                                TRX SERV   28
```


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 APARs

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

HiperDispatch

■ OA36054

- Beginning with z/OS 1.13 when running on an IBM zEnterprise z196 or z114 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

Workload Promotion

■ OA30068

- PDSE hang can occur on various PDSE latches due to address space getting swapped out by WLM while holding PDSE resources
- PDSE contention couldn't be resolved by blocked workload support since latch holder was swapped out
 - Would require the address space to be made non-swappable
- PDSE latch processing is changed to add SYSEVENT ENQHOLD function to allow SRM to boost the service of the latch holder
 - Improves swapin recommendation value

■ OA35373

- Unilaterally swapped address spaces are not being exchanged swapped for long periods of time. Max swap out time is set to the minimum of OUCBOUTT or 30 mins
- This is too long to leave work swapped out

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

Local Lock Promotion - LCK

- 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

DB2 Distributed (DDF) and z/OS

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

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

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
 - Often the vast majority of the DB2 9 'non-preemptable' SRB processing in DBM1
 - Will depend on workload mix

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

Calculating zIIP Offload

- For the workload changed to pre-emptible SRBs 100% of the CPU time is zIIP eligible

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

- Calculate the potential zIIP busy by:

$$\text{NonPreempt SRB Time} / \text{Interval Time} = \% \text{ of a CP Eligible for zIIP}$$

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%

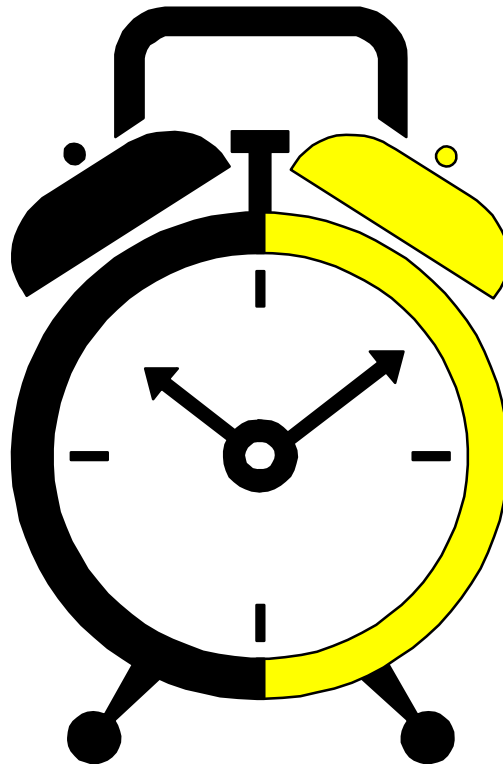
Example of Potential Offload

- Eight Data Sharing Members during Batch 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
0.57	0.42	0.45	0.42	73%
1.04	0.71	0.76	0.71	68%
0.58	0.36	0.38	0.36	63%
0.28	0.22	0.23	0.22	81%
0.23	0.14	0.21	0.14	62%
0.12	0.10	0.10	0.10	82%
0.62	0.44	0.46	0.44	70%
0.52	0.24	0.41	0.24	45%

Addendum

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



TD105930 - Hiperdispatch and SAP DB Servers

Typically, customers run SAP DB Servers in their own z/OS LPARs and this leads to unique recommendations for Hiperdispatch for these LPARs

- z/OS R13: Use HiperDispatch (z10 and z196)
- z/OS R12 or earlier: Depends on the hardware configuration
 - z196 – Enable HiperDispatch
 - Main goal is to cluster each LPAR's work to a limited number of chips in the same book as this provides hardware cache benefits.
 - Installations with a very large LPAR processing SAP work may need to do more analysis on whether or not HiperDispatch should be used.
 - A "large" installation would be an LPAR which has more than five logical zIIPs defined AND the logical zIIPs are more than 80% busy.
 - For environments which meet this definition of "large", the amount of IIPCP time incurred in the environment may warrant running the environment with HiperDispatch disabled
 - z10 - Disable HiperDispatch
 - With the z10 the main goal is to cluster each LPAR's work to a specific book
 - For an SAP environment where the LPAR fits on one book, the benefits of HiperDispatch in an SAP environment are nominal
 - If the SAP LPAR is on a processor with purchased CPs (GCPs, zIIPs) on more than one book, additional analysis may be needed with respect to the HiperDispatch recommendation

z/OS 1.13 Performance Line Item Review

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

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

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

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

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

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

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

```

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

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

SYSTEM	RESPONSE TIME EX VEL%	PERF INDX	AVG ADRSP	--EXEC USING%--				EXEC DELAYS %			-USING%-		--- DELAY % ---				
				CPU	AAP	IIP	I/O	TOT			CRY	CNT	UNK	IDL	CRY	CNT	QUI
*ALL	--N/A--	88.2	0.2	47.0	0.0	0.0	0.0	0.2	0.0		0.0	0.0	38	62	0.0	0.0	0.0
SYSD		88.2	0.2	15.0	0.0	0.0	0.0	0.3	0.0		0.0	0.0	40	60	0.0	0.0	0.0
SYSE		88.6	0.2	17.0	0.0	0.0	0.0	0.3	0.0		0.0	0.0	35	64	0.0	0.0	0.0

-----RESPONSE TIME DISTRIBUTIONS-----

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

z/OS 1.13 RMF GRS & Supervisor Delay Monitoring

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

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 1.13 SMF Logstream Support

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

z/OS 1.12 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)
<http://www-03.ibm.com/support/techdocs/atmastr.nsf/WebIndex/TD103286>

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
    
```