z/OS Workload Management

Fall 2017 Update for z14 and z/OS V2.3

Horst Sinram, STSM, z/OS Workload and Capacity Management, sinram@de.ibm.com
Andreas Henicke, Developer, z/OS Workload Management, andreas.henicke@de.ibm.com
IBM Germany Research & Development July 14th 2017

IBM

Trademarks

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

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

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

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

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

The following are trademarks or registered trademarks of other companies.

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

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

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

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 SE 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

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview

z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Two new capping options

z/OS release	V2.2	V2.1
Function		
z13 GA2 LPAR Absolute group capping	*OA47752*	*OA47752*
Absolute MSU capping	*OA49201*	*OA49201*

- **LPAR absolute group capping**
  - Feature of PR/SM as of z13 GA2, and z13s

  - Like LPAR absolute capping but for a group of LPARs

  - Specified on the HMC as number of processors with 2 decimal places (like 3.75). All processor classes supported.
  - Recognized by WLM as possible limit to the LPAR capacity

- **WLM Absolute MSU capping**
  - Function of WLM provided by APAR OA49201. Requires zEC12 GA2 or later.

  - Similar to WLM defined capacity or group capacity but LPAR will always be capped
    - Independent of 4 hour rolling average consumption.
    - General purpose processor
  - Specified in IEAOPTxx. Limit is the LPAR defined capacity or group capacity specified on the HMC **in MSU**.

6

Using absolute MSU capping

IEAOPTxx ABSMSUCAPPING=	
<u>NO</u>	Defined capacity limits and group capacity limits should be enforced only while the long term four hour rolling average consumption exceeds the respective limit (existing and usually desired behavior).
YES	Defined capacity limits and group capacity limit should be enforced **permanently, independently of the long term four hour rolling average consumption**. Becomes effective on zEC12 GA2 or later.

- AbsMSUcapping=Yes limits LPAR consumption to a certain MSU number at all times.
  - I.e., the system loses the flexibility of consuming above the defined capacity limit while the four hour rolling average is below the limit.

- Limit remains stable even when CEC configuration changes, e.g. through On/Off CoD or CBU activations or deactivations.
  - Absolute MSU capping is an effective means to permanently limit the consumption of an LPAR to a specific MSU figure at all times
    - including times when the *four-hour rolling average* does not exceed the defined limit.

7

Using absolute MSU capping with group capacity

- When used with an LPAR capacity group:
    - Limit on behalf of the group entitlement will always be enforced
        - Regardless of the *four-hour rolling group average* consumption.

    - As with AbsMSUcapping=NO, an LPAR is allowed to take benefit of the unused group capacity
        - Unless the LPAR is also capped via other LPAR limits.

    - All members of a capacity group that use AbsMSUcapping=YES will permanently enforce the limit on behalf of the capacity group.

    - All members of a capacity group that do *not* use AbsMSUcapping=YES will be capped while the group *four-hour rolling group average* consumption is greater or equal to the group limit

Agenda

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview

z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Comparison of capping types

➡️ New capping types (spring 2016)

Type of capping	Scope	Specification unit	Proc types	Stability of limit under configuration changes	Suitable to isolate LPARs or LPAR groups	Control point
Initial (hard) capping	LPAR	LPAR share of CPC capacity	Any	-	+	SE/HMC
LPAR Absolute capping (zEC12 GA2 and later)	LPAR	Fractional #processors		O	+	
➡️ LPAR Group Absolute Capping (z13 GA2 and later)	Group of LPARs	Fractional #processors		O	+	
Defined capacity (DC, soft capping)	LPAR	MSU (4HRA)	CP	+	-	
LPAR group capacity (GC, soft capping)	Group of LPARs	MSU (4HRA)		+	-	
➡️ Absolute MSU Capping	LPAR or Group	MSU		+	+ (CP only)	SE/HMC + IEAOPT
Resource group capping	Groups of service classes in Sysplex or per LPAR	Unweighted CPU SU/sec, fraction of LPAR share, or fractional #CPs	CP*	+	N/A	WLM Policy
Logical configuration	LPAR	Integer #processors	Any	O	+ but coarse grain	HMC+OS

🟧 PR/SM controlled   🟨 WLM controlled, PR/SM enforced   🟦 WLM controlled

10

© Copyright IBM Corp. 2017

Which capping techniques may be combined?

Type of capping→	Initial (hard capping)	LPAR Absolute capping	LPAR Absolute group capping	Defined capacity [1]	LPAR group capacity [1]	Resource group capping
Initial (hard capping)		+	+	-	-	+
LPAR Absolute capping	+		+	+	+	+
LPAR Group Absolute capping	+	+		+	(2)	+
Defined capacity [1]	-	+	+		+	+
LPAR group capacity [1]	-	+	(2)	+		+
Resource group capping	+	+	+	+	+	

(1) Includes ABSMSUCAPPING=NO and ABSMSUCAPPING=YES
(2) Is compatible if LPAR Group is the same

Reporting enhancements for CICS, IMS, and Mobile and Cloud Workloads

- **Mobile Workload Pricing** (MWP) is an IBM Software Pricing option, announced in May 2014

- **Workload Pricing for Cloud** (zWPC) is an IBM Software Pricing option, announced in July 2016

- For eligible software both can reduce the cost of transactions that originate from mobile devices or new public cloud workloads
  - MWP and zWPC can mitigate the impact of such workloads on sub-capacity license charges, specifically in the cases where higher mobile or cloud transaction volumes may cause a spike in machine utilization

- Reporting enhancements for CICS, IMS, and Mobile and Cloud Workloads introduce WLM enhancements that can simplify the identifying and reporting of the mobile- or cloud-sourced transactions and their processor consumption

Enablement of Reporting Enhancements for Mobile Workloads

z/OS release / Function	z/OS V2.1	z/OS V2.2	Other
WLM: Reporting Enhancements for Mobile Workloads	OA47042	OA47042	
RMF Reporting Enhancements for Mobile Workloads	OA48466	OA48466	
z/OSMF Reporting Enhancements for Mobile Workloads	PI47638	PI47638	
CICS TS Reporting Enhancements for Mobile Workloads			CICS 5.3
IMS TM Reporting Enhancements for Mobile Workloads			IMS 14 PI46933 (available) PI51948 (1H2016*)
SCRT and Billing System Support			SCRT 23.13.0 (mobile, available) SCRT 24.10.0 (cloud, available)

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.
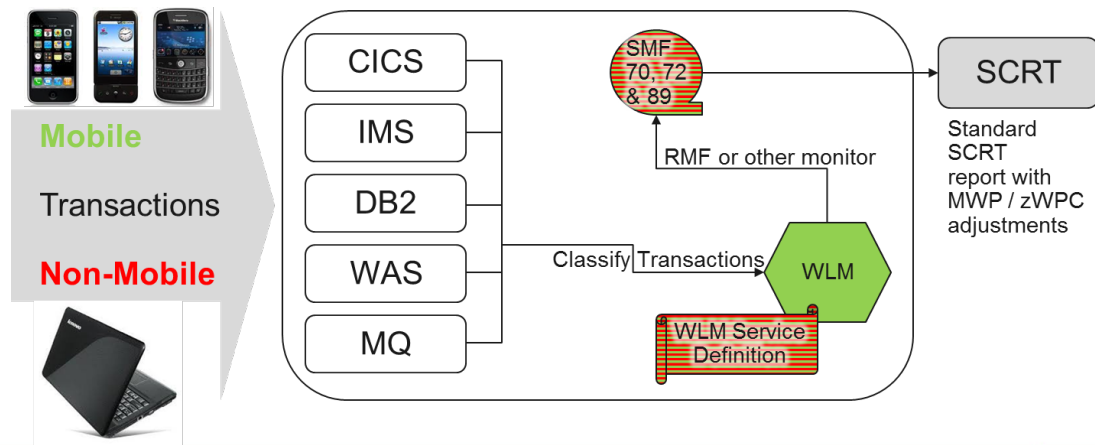
Identifying Mobile and Cloud Workload

MWP and zWPC offer a discount on MSUs consumed by transactions that originated from a mobile device or new public cloud workloads

To take advantage of this discount, you need a process, agreed upon by you and IBM, to **identify (tag and track)** mobile- or cloud-sourced transactions and report on their consumption

→ NEW Identify mobile or cloud transaction via a **transaction level attribute** in the WLM service definition

- Processor consumption data aggregated by WLM
- Reporting integrated into standard performance monitors (RMF) and low volume SMF records
- Applicable to wide range of workloads, including enclave work and CICS/IMS work

Mobile

Transactions

Non-Mobile

CICS

IMS

DB2

WAS

MQ

Classify Transactions

WLM

SMF 70, 72 & 89

RMF or other monitor

WLM Service Definition

SCRT

Standard SCRT report with MWP / zWPC adjustments

Identifying Mobile and Cloud Workload (cont.)

- In your WLM classification rules, classify transactions as **MOBILE, CATEGORYA or CATEGORYB**
- The assigned attribute is independent from the assigned service and report class
  - Eliminates the need for using new dedicated classes for mobile or cloud workload reporting
- The assigned attribute is transparent to subsystems
- WLM tracks and reports the total and the **MOBILE, CATEGORYA and CATEGORYB CPU consumption** for all service and report classes
  - **With exploiting levels of CICS and IMS, CPU consumption data is also available for CICS and IMS transaction service and report classes that previously did not report any CPU consumption data**
  - **Subsystems using independent enclaves can participate transparently; only the classification rules need to be updated.**
- WLM also aggregates and reports the **system-wide MOBILE, CATEGORYA and CATEGORYB consumption** data

Using the New Support – An Example

Scroll right twice

Using the New Support – RMF Sample Workload Activity Report (I)

Get processor consumption data for CICS and IMS transactions from **low-volume** SMF records:

```
REPORT BY: POLICY=BASE          WORKLOAD=CICSWKLD       SERVICE CLASS=CICSFAST     RESOURCE GROUP=*NONE
                                                        CRITICAL      =NONE
                                                        DESCRIPTION   =CICS fast Service class


   -TRANSACTIONS-    TRANS-TIME HHH.MM.SS.TTT
   AVG        0.00   ACTUAL                 1.432
   MPL        0.00   EXECUTION                710
   ENDED      632    QUEUED                     0
   END/S     10.58   R/S AFFIN                  0
   #SWAPS       0    INELIGIBLE                 0
   EXCTD      631    CONVERSION                 0
   AVG ENC    0.00   STD DEV                    0
   REM ENC    0.00
   MS ENC     0.00


                            MOBILE :  CP   99.16   AAP/IIP ON CP   0.00    AAP/IIP    0.00
```

Using the New Support – RMF Sample Workload Activity Report (II)

Use processor consumption data for CICS and IMS transactions for **charge-back**:

```
--------------------------------------------------------------------------------
REPORT BY              WORKLOAD=CICSWKLD    SERVICE CLASS=CICSFAST    RESOURCE GROUP=*NONE
                                            CRITICAL    =NONE
                                            DESCRIPTION =CICS fast Service class
...
   TRANSACTIONS APPL% :   TOTAL :   CP 185.92   AAP/IIP ON CP   0.00   AAP/IIP   0.00
                          MOBILE :  CP  99.16   AAP/IIP ON CP   0.00   AAP/IIP   0.00
--------------------------------------------------------------------------------
REPORT                                      REPORT CLASS=BTRAD
                                            DESCRIPTION =Report Class for tradit.banking
...
   TRANSACTIO        :   TOTAL     CP  74.32   AAP/IIP ON CP   0.00   AAP/IIP   0.00
                         MOBILE :  CP  24.84   AAP/IIP ON CP   0.00   AAP/IIP   0.00
--------------------------------------------------------------------------------
   REPORT B                                  REPORT CLASS=BDIRECT
                                            DESCRIPTION =Report Class for direct banking
...
   TRANS          L :   CP 111.59   AAP/IIP ON CP   0.00   AAP/IIP   0.00
                  MOBILE :  CP  74.32   AAP/IIP ON CP   0.00   AAP/IIP   0.00
```

Overall consumption for CICS transaction service class

40% can be charged back to the traditional banking division

60% can be charged back to the direct banking division

Using the New Support – RMF Sample Workload Activity Report (III)

Check **region overhead** of CICS and IMS regions:

```
REPORT BY: POLICY=BASE        WORKLOAD=STCWORK      SERVICE CLASS=VEL80      RESOURCE GROUP=*NONE
                                              CRITICAL    =NONE
                                              DESCRIPTION =Velocity=80 goal

  -TRANSACTIONS-     TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---    SERVICE TIME   ---APPL %---   --PROMOTED--  ----STORAGE----
  AVG      3.00      ACTUAL              0     SSCHRT   0.0  IOC        0     CPU 132.597    CP    221.98   BLK   0.000   AVG    2481.98
  MPL      3.00      EXECUTION           0     RESP     0.0  CPU     7266K    SRB   0.001    AAPCP   0.00   ENQ   0.000   TOTAL  7445.94
  ENDED       0      QUEUED              0     CONN     0.0  MSO   180672K    RCT   0.000    IIPCP   0.00   CRM   0.000   SHARED    0.00
  END/S    0.00      R/S AFFIN           0     DISC     0.0  SRB      224     IIT   0.000                  LCK   0.020
  #SWAPS      0      INELIGIBLE          0     Q+PEND   0.0  TOT   187938K    HST   0.000    AAP    0.00   SUP   0.000   -PAGE-IN RATES-
  EXCTD       0      CONVERSION          0     IOSQ     0.0  /SEC    3146K    AAP   0.000    IIP    0.00                 SINGLE     0.0
  AVG ENC  0.00      STD DEV             0                                   IIP   0.000                                BLOCK      0.0
  REM ENC  0.00                                            ABSRPTN  2321K                                               SHARED     0.0
  MS ENC   0.00                                            TRX SERV 2321K                                               HSP        0.0


  TRANSACTIONS APPL% :  TOTAL :  CP  36.06  AAP/IIP ON CP   0.00   AAP/IIP   0.00
                        MOBILE :  CP   0.00  AAP/IIP ON CP   0.00   AAP/IIP   0.00

  -------------------------------------------------SERVICE CLASSES BEING SERVED-------------------------------------------------
  CICSFAST
```

Region-only service, excluding transaction service = the region overhead

Overall region consumption including transaction service

Using the New Support – System-wide Consumption

- WLM aggregates the system-wide total, MOBILE, CATEGORYA and CATEGORYB processor consumption data, accumulates the values into a rolling 4-hour average, and reports it to performance monitors
- From there, RMF or other monitors are be able to pick it up and report into SMF70

Rolling 4-Hour Average with MWP Adjustment
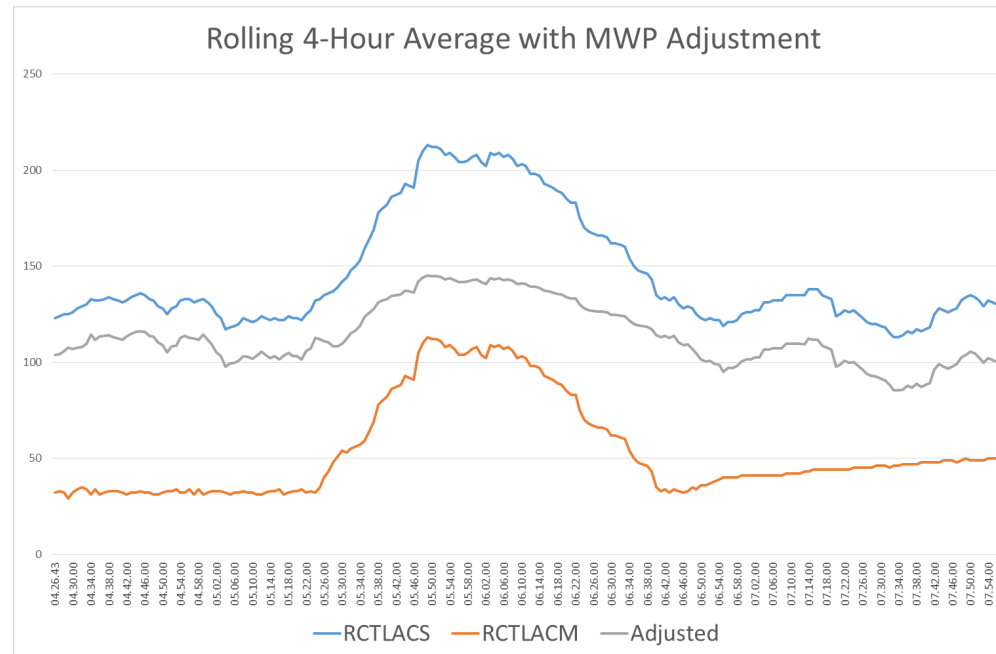
XML Format WLM service definitions recommended

- For several releases WLM has supported to store a service definitions in **XML format**
  - z/OSMF WLM task
  - ISPF Administrative Application: "Save as XML"…

- XML format avoids particular problems with the ISPF tables format, namely coexistence behavior, when a new functionality level needs to be introduced, and the number of table columns needs to be extended.
  - **For example, OA47042 introduces such a change**
  - **And there is more to come!**

- **Recommendation:
  Use the XML-format for your WLM service definition data sets**

- In the next release, XML-format will be the default for the "Save" and "Save as" actions in the ISPF Administrative Application*

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Agenda

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview

z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Why new z/OS service definition objects?

Cloud workload paradigm asks for new ways of metering workloads in multi-tenant environments

IBM Z business asks for an infrastructure to support novel pricing options

- A Tenant Report Class (TRC) is similar to a WLM Report Class.
  TRCs are assigned through WLM classification and are always associated with a Tenant Resource Group.

- A Tenant Resource Group (TRG) is somewhat similar to a WLM Resource Group and can be associated with tenants or solutions.

- TRGs aggregate consumption data and can optionally be used to apply consumption limits.

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Overview of WLM/SRM Enhancements for Container Pricing for IBMZ

- New panels for defining and modifying Tenant Resource Groups

- New panels for defining and modifying Tenant Report Classes

- Resource Group (RG) and Tenant Resource Group (TRG) enhancements
  - New Sysplex-wide Type 4 limit expressed at a scale of "MSU"
  - Optionally, specialty processor consumption can be counted towards the limit

- A new service definition option that allows to disable "Discretionary Goal Management" globally

- WLM/SRM programming services are enhanced
  - A new IWM4QTNT service allows monitoring products to retrieve TRG-level consumption data
  - Various APIs provide TRC and TRG indications

- *WLM TRG level data will be reported by monitoring products, such as RMF*
  - *Also in SMF type 70 record*

- *SMF70 TRG data can be consumed in SCRT*

- *Various other components to be updates (SMF, SDSF, z/OSMF, …)*

TRG and TRC Definition*

- The WLM Administrative Application Level increased to 32.

- Tenant Resource Groups and Tenant Report Classes can be defined via new menu items.

- Specification of these new objects will increase the functionality level of the service definition to 32.

```
Functionality LEVEL032      Definition Menu      WLM Appl LEVEL032

------------------------------------------------------------

Definition data set  . . : 'WLM.DEMO.SRVDEF.XML'
Definition name  . . . . . PROD01    (Required)
Description  . . . . . . . Production service definition


Select one of the following options.

__      1.  Policies                  12. Tenant Resource Groups
        2.  Workloads                 13. Tenant Report Classes
        3.  Resource Groups
        4.  Service Classes
        5.  Classification Groups
        6.  Classification Rules
        7.  Report Classes
        8.  Service Coefficients/Options
        9.  Application Environments
        10. Scheduling Environments
        11. Guest Platform Mgmt Provider
```

Tenant Resource Group (TRG) Definition*

- The TRG name is mandatory (8 char)

- Description, Tenant ID, Tenant Name are optional and are expected to be used in a z/OS cloud context

- For qualified offerings, a 64 char Solution ID needs to be provided.

- Exactly enter (paste) the IBM provided Solution ID string

- WLM performs sanity check only.

- Solution ID considered during SCRT processing. Multiple TRGs may specify same Solution ID

- TRG capacity limits should not be specified unless there is a need to limit processor consumption. There is NO minimum consumption limit.

```
                 Create a Tenant Resource Group

Enter or change the following information:
Tenant Resource Group Name  TRGDEM01  (required)

Description . . . . . . . . Sample TRG

Solution ID:
 ZS1ZZZA-I1F9FAD-6B20999D15-85370C8D-726F-458C-B84A-37D01D-A10E44

Tenant ID . . . . . . . . . _____

Tenant Name . . . . . . . . Solution Newapp

Define Capacity: __
             1.  In Service Units           (Sysplex Scope)
             2.  As Percentage of the LPAR share (System Scope)
             3.  As a Number of CPs times 100  (System Scope)
             4.  In accounted workload MSU    (Sysplex Scope)

Maximum Capacity  . . . . . . . . . . . . _____

Include Specialty Processor Consumption NO      (YES or NO)
```

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Tenant Report Class (TRC) Definition*

- The TRC name is mandatory (8 char)
  - Name must be unique (also across report classes)
  - Up to 2047 Report Classes and Tenant Report Classes can be defined

- The TRG name is required, i.e. any TRC must be associated with a TRG

- Monitoring interfaces and monitors report on TRC as on standard report classes

```
                 Create a Tenant Report Class

-----------------------------------------------------

Enter or change the following information:


Tenant Report Class Name . . . T_CDC      (Required)
Description  . . . . . . . . .  TRC including CDC region


Tenant Resource Group Name . . TRGDEM01  (Required; name or ?)
```

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Comparision of TRG and RG Capping Types*

Type 1	Type 2	Type 3	Type 4
Raw CPU+SRB service units ("Raw" meaning that Service Definition Coefficients are not applied)	Percent of CP LPAR share (even if specialty processor consumption included). May exceed 100%.	Percent of one CP processor (even if specialty processor consumption included)	Processor consumption expressed in "accounted workload MSU"
Limit applies to Sysplex	Limit applies to each System	Limit applies to each system	Limit applies to Sysplex

For **all** (T)RG types only captured TCB and SRB times are counted towards the limit. The limit is enforced based on a one minute average (i.e., no 4HRA). Up to 32 RGs **plus 32 TRGs** may be defined.
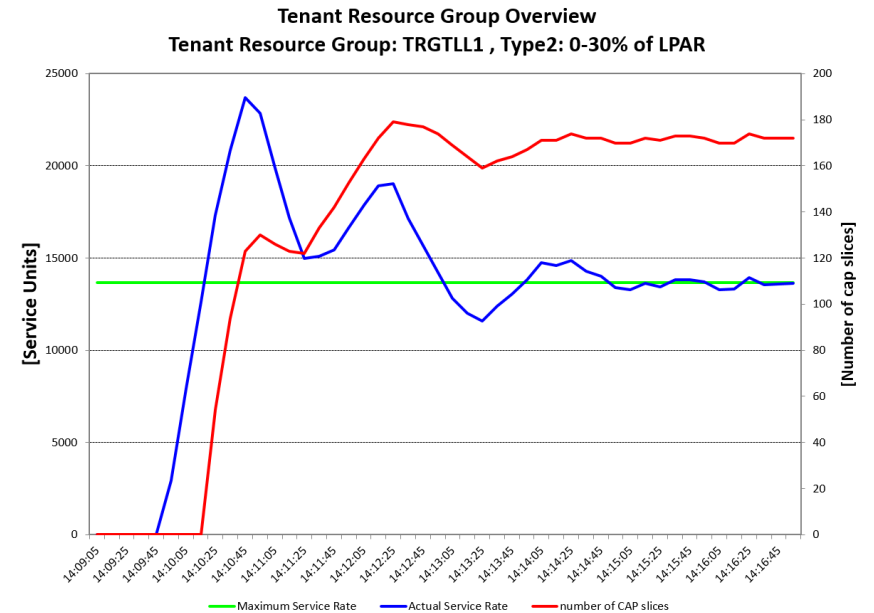
* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

TRG Capping*

▪ TRG capping is based on Resource Group capping and inherits its characteristics.

• Time is divided into 256 "slices".
  In any slice the whole (T)RG can be set
  -dispatchable (called awake slice)
  -non-dispatchable (cap slice)

• The cap pattern is adjusted every 10 sec based on the average of the last minute

• (T)RG consumption will vary base on demand, mix of dispatch priorities, number of dispatchable units and number of processors.
  • The system will attempt to over-cap the work, i.e., the consumption will be throttled to remain below the limit. Depending on the characteristics this may not always be possible.
  • Usually consumption levels in within minutes

• Very latency sensitive work not a good candidate for capping. Multiple TRGs with same Solution ID may be used when needed.

Sample cap pattern showing work that is capped 50% of time.

Tenant Resource Group Overview
Tenant Resource Group: TRGTLL1 , Type2: 0-30% of LPAR

Maximum Service Rate   Actual Service Rate   number of CAP slices

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

New IWMW4QTNT Service*

- Returns TRG related data in the return area mapped by IWMWQTAA.

- TRG description data

- Long term average (4HRA) of CP consumption in MSU.
  The captured TRG consumption is extrapolated to the LPAR consumption, i.e., apportioned uncaptured time is included.

- Consumption in Service Units (SU) on each processor type

```
IWM4QTNT    ANSAREA=xansarea
            ,ANSLEN=xanslen
            ,QUERYLEN=xquerylen
            ,ANSTOKN=xanstokn
            [,RETCODE=xretcode]
            [,RSNCODE=xrsncode]
            [,PLISTVER={xplistver|IMPLIED_VERSION}]
            [,MF=S]
            [,MF=(L,xmfctrl{,xmfattr|0D})]
            [,MF=(E,xmfctrl{,COMPLETE})]
```

For each TRG, IWM4QTNT returns increasing aggregated values.
Monitoring products can derive interval values by computing deltas, e.g. for reporting in SMF type 70.1 records

New service definition option to deactivate DGM

Discretionary Goal Management (DGM) can improve throughput of discretionary work by throttling certain vastly over-achieving non-discretionary work.

It "throttles" through dynamically created resource groups.

With this support, DGM can optionally be disabled.

- Service Coefficient/Service Definition Options

- Enter or change the Service Coefficients:

- CPU 1.0   (0.1-99.9)
  IOC 0.5   (0.0-99.9)
  MSO 0.0000   (0.0000-99.9999)
  SRB 1.0   (0.0-99.9)

- Enter or change the service definition options:

- I/O priority management . . . . NO   (Yes or No)
  Enable I/O priority groups . . . . NO   (Yes or No)
  Dynamic alias tuning management . . . . . NO   (Yes or No)
  Deactivate Discretionary Goal Management   Yes (Yes or No)

Agenda

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview

z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Enablement of More granular resource controls*

z/OS Release / Function	z/OS V2.2	z/OS V2.1
WLM/SRM Support	OA50845	OA50845
RMF Reporting Enhancements	OA50760	OA50760
z/OS Supervisor Support	OA50953	OA50953
z/OS RSM (Real Storage Manager)	OA51171	OA51171
z/OSMF Reporting Enhancements	PI71118	PI71084

More granular resource controls*

- **Purpose:** provide more granular control over CPU and memory consumption by workload

- Initial focus on demanding workloads that run only on specialty engines like Java batch, SPARK and other analytics, and zCloud workloads

- New controls:
  - ➢ Honor Priority by service class
  - ➢ Memory Limit for resource groups

  - More to come ☺

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

More granular resource controls*
Honor Priority by Service Class

```
---------------------------------------------------------------------------
                      Modify a Service Class                  Row 1 to 4 of 4
Command ===>  _____

Service Class Name . . . . . . :  FAST
Description  . . . . . . . . . .  Velocity=80 goal_____
Workload Name  . . . . . . . . .  STCWORK_   (name or ?)
Base Resource Group  . . . . . .  _____   (name or ?)
Cpu Critical . . . . . . . . . .  NO_        (YES or NO)
I/O Priority Group . . . . . . .  NORMAL_    (NORMAL or HIGH)
Honor Priority . . . . . . . . .  NO_____   (DEFAULT or NO)

Specify BASE GOAL information. Action Codes: I=Insert new period,
E=Edit period, D=Delete period.

        -- Period --  -----------------  Goal -----------------
Action  #  Duration   Imp.  Description
  __    1  2000000      2   Average response time of 00:00:01.000
  __    2  2000000      3   Average response time of 00:00:10.000
  __    3  _____   4   Execution velocity of 80
*********************************** Bottom of data ***************************
```

- Specifies whether work in this service class is exempted from default IFAHONORPRIORITY and IIPHONORPRIORITY processing

- Also for Service Class Overrides

- Limitation to specialty engines enforced collaboratively with SUP

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Recommendations for HonorPriority

- Some zIIP work may be very latency sensitive and require to be dispatched quickly .
  - Namely some DB2 work, such as prefetch SRBs.
  - zIIP capacity may be constrained but CP capacity might be available to help

- Recommendation:

  - At the system level (IEAOPTxx) specify or default to IIPHonorPriority=Yes to allow CPs to help zIIP work.

  - Use the service class specific HonorPriority=No to selectively exclude work from receiving help.
    – Examples could be SPARK or Java batch that you do not want to be processed on general purpose processors

More granular resource controls*
Memory Limit for Resource Groups

```
-----------------------------------------------------------------------------
                        Modify a Resource Group
Command ===> _____

Enter or change the following information:

Resource Group Name  . . . . . : MEMLIMIT
Description  . . . . . . . . . . Provide a memory limit for work

Define Capacity:
__      1.   In Service Units (Sysplex Scope)
        2.   As Percentage of the LPAR share (System Scope)
        3.   As a Number of CPs times 100 (System Scope)
Minimum Capacity . . . . . . .  _____
Maximum Capacity . . . . . . .  _____

Memory Limit (System Scope)      24_____     GB
```

- Specifies the maximum amount of memory that address spaces associated with the resource group through classification may consume on the local system (System Scope)
- The attribute is specified as absolute value in GB in the range 1 – 99,999,999.
- Also for Resource Group Overrides
- Memory limit enforced collaboratively with RSM

How do memory pools work?

- A Resource Group definition, as part of the WLM service definition, has Sysplex scope.

- All address spaces get classified by WLM and get a Service Class assigned
  - **If** the assigned Service Class is associated to a resource group **and** the resource group specifies a memory limit, WLM notifies RSM to connect the address space to a memory pool

- When the first address space connects to a memory pool, RSM creates that pool. The pool name is equal to the resource group name.

- An RSM memory pool represents only a logical pool (=upper limit).
  It is **not** dedicating or reserving real storage.
  → A memory pool can specify a size exceeding the real storage of a given system

- All work in the system is managed towards the global pool (total real storage).
  Address spaces connected to a memory pool are also subject to the pool limit.
  - When the pool limit is approached self-stealing is initiated to keep the number of frames within the limit.

- Address spaces may be temporarily deferred, if the pool limit would be exceeded by adding the space.

Memory pools - current limitations

- When the memory pool support becomes available the following limitations apply:

  - An active address space cannot be reclassified to another defined memory pool.
    - The address space has to terminate in order to be reclassified, or it must be reclassified to the global pool and then reclassified again to the new pool.
    - Exception: initiator address spaces for a new job

  - Resource Group memory pool limits cannot be *decreased* while it is defined into a policy.
    - The only way to decrease the limit is to activate a policy that does not have the resource group defined and then activate a policy that defines the pool with a smaller limit.
    - However, a pool can be dynamically *increased* via a new policy activation.

  - Memory related Sysevents (such as STGTEST) are not memory pool aware

  ➔ IBM recommends that you use memory pools when it is required to limit memory consumption for new workloads such as Apache SPARK that provided guidance on how to operate them in a memory pool.

New WLM/SRM Messages

- **IRA450E MEMORY POOL** *mmmmmmmm* **MEMORY LIMIT REACHED -** *nnn* **JOBS DEFERRED**
  – The system detected a memory pool shortage in memory pool *mmmmmmmm*. Jobs which start and associated to this memory pool are deferred until the memory pool shortage is resolved.

- **IRA452I MEMORY POOL** *mmmmmmmm* **BELOW THE MEMORY LIMIT**
  – There is no longer a memory pool shortage for memory pool *mmmmmmmm.*

- **IRA459I MEMORY POOL -** *text*
  – A new WLM policy was activated or an address space was reclassified to a new WLM service class via the RESET command. Possible problems:
  – **MEMORY LIMIT REDUCTION NOT SUPPORTED**
  – **RECLASSIFY TO MEMORY POOL** *mmmmmmmm* **IS NOT SUPPORTED**
  – **CREATION PROBLEM mmmmmmmm**
  – **INTERNAL PROBLEM mmmmmmmm RSN=xxxx**

Agenda

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview

z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

WLM/SRM Exploitation of IBM z14

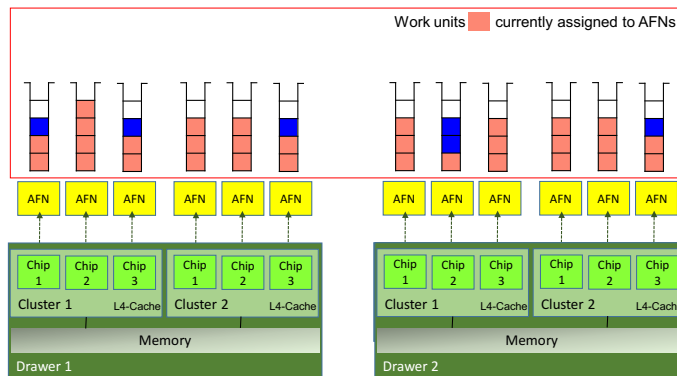| z/OS release / Function	V2.3	V2.2	V2.1
WLM/SRM Support for IBM z14	**+**	**OA50144**	**OA50144**
SRM Support for RSM Performance	**+**	**OA51904**	**OA51904**

OA50144: Hiperdispatch Memory Affinity

- HiperDispatch (HD) implements a smarter HD balancer move algorithm for 'high storage consumers' (HSC) which considers now the processor topology of the HD affinity nodes (AFN) and thus implicitly of memory affinity
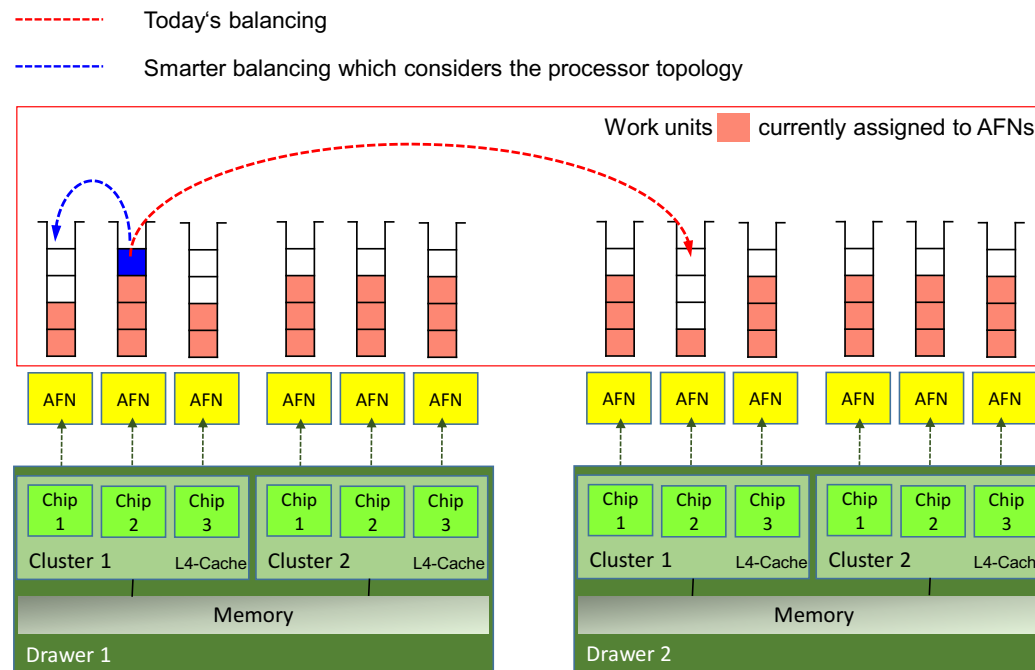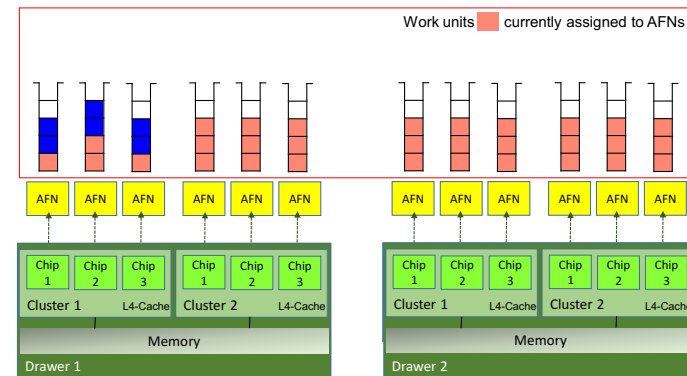
OA50144: HiperDispatch Topology Enhancements

- HiperDispatch (HD) implements a smarter HD address space breakup algorithm which considers the processor topology of the broken up address space and its split off TCBs

OA51904: SRM support for RSM performance items

- z/OS Freemained Frame management is extented to high frames, sometimes referred to as authorized storage

- Adaptive region Freemained Frame Target (FFTarget) calculation based upon requirements (1% to 40%)

- Freemained Frames now get stolen by SRM based upon importance (from unimportant address spaces first)
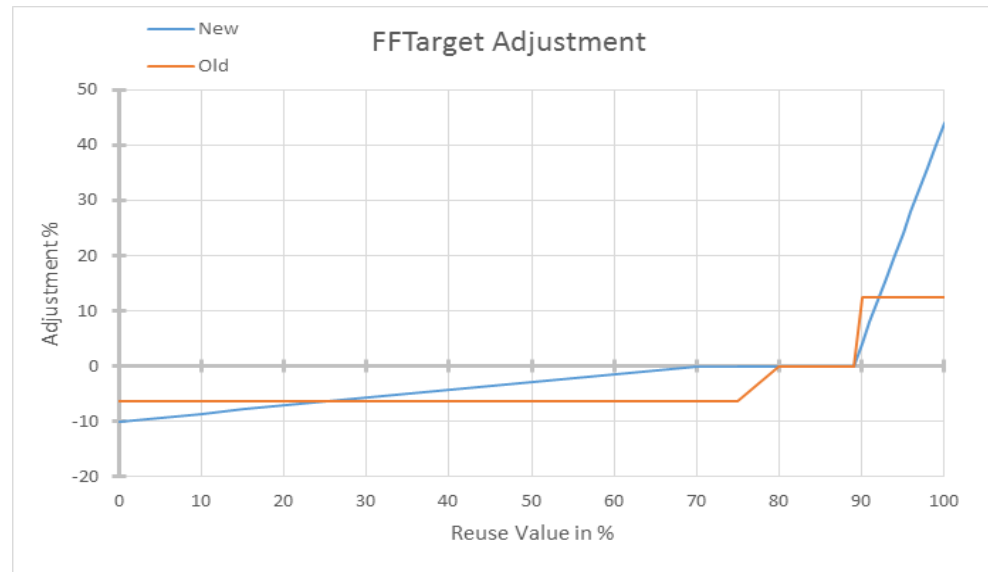
Agenda

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview

z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

WLM/SRM support overview for IBM z13

z/OS release Function	V2.2	V2.1	V1.13
z13 GA2 LPAR Absolute group capping	OA47752	OA47752	
z13 Support (base)	+	OA43622 OA47021	OA43622
z13 HiperDispatch Optimizations	OA47968 (Included in GA code)	OA47968	OA47968
zIIP SMT Support	+	OA43622	
Hiper-Dispatch z13 & zEC12 — Unpark while capped / Unused capacity refinement / Prime cycle elimination	+	OA43622	
SRM storage management changes in support of RSM for z13	OA48858	OA44504 OA46396 OA48858	OA44504 OA46396 OA48858

Base z13 support

- New limits for z13
  - 85 LPARs
    - Up to 141 processors per CPC
    - Up to 141-way on z/OS V2.1 (non-SMT mode)
    - Up to 128-way on z/OS V2.1 (SMT mode), or z/OS <V2.1
    - Maximum active threads in SMT mode is 213 with zIIP:CP ratio of 2:1
- New Cache topology
  - Chip, node, drawer
  - No longer using "books"
  - z/OS HiperDispatch uses new topology information to place work topologically close – to maximize cache efficiency
  - New as-is tool available for download from the WLM homepage WLM Topology Report
- No zAAPs
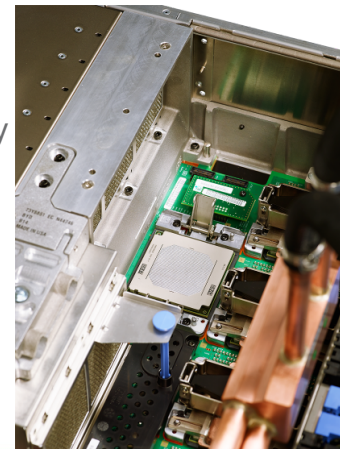
zIIP SMT support

- With z13, Simultaneous multithreading (SMT) allows up to two instructions streams per core to run simultaneously to get better overall throughput and better utilize the processor hardware units

- On z/OS, SMT is available for zIIP processing:

  - Two concurrent threads are available per core

  - Capacity (throughput) usually increases

  - Performance may be superior using single threading
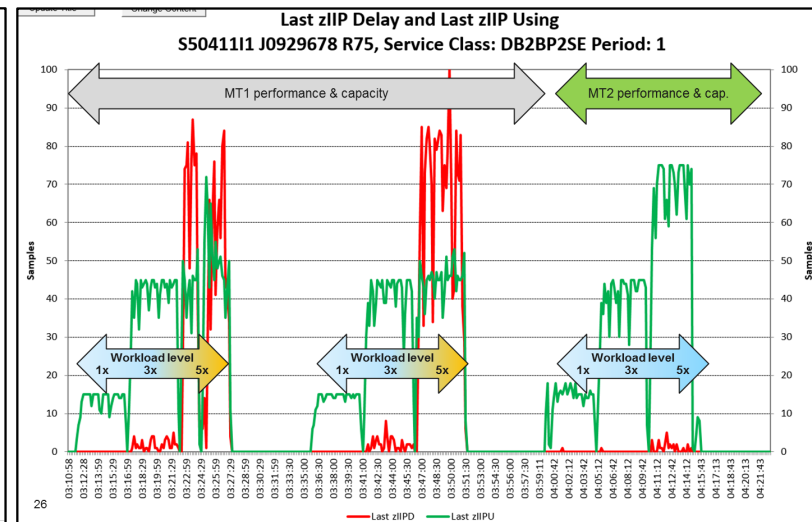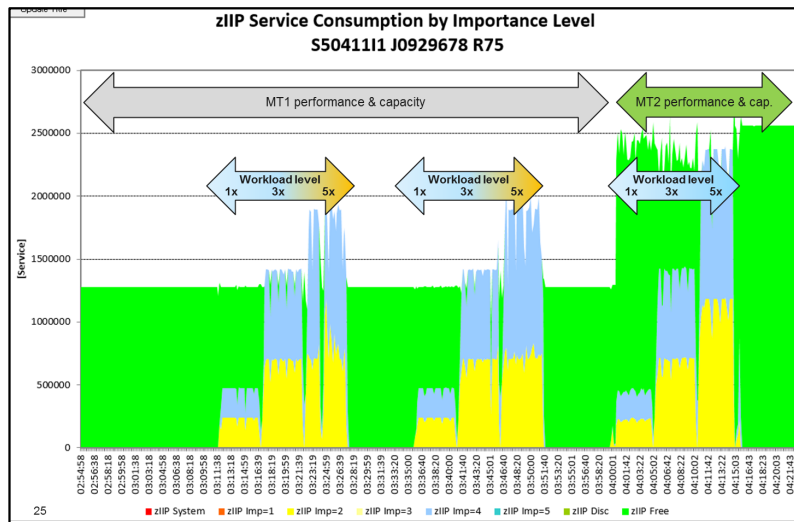
At a glance:
What is new with multithreading support?

New support

LOADxx: **PROCVIEW CPU**	LOADxx: **PROCVIEW CORE**	
• This the existing mode that you are used to • z/OS does only know processors (CPUs)	• z/OS is core-aware and maps *CPUs* to *cores*. • Core syntax for CPU/core related commands.	

IPL required

MT1 Mode	**MT2 Mode**
IEAOPTxx: MT_ZIIP_MODE = 1	IEAOPTxx: MT_ZIIP_MODE = 2

SET OPT=yy

Possible on any supported hardware | z13 only

MT1 performance & capacity | MT2 performance & cap.

HiperDispatch=No possible | HiperDispatch=Yes enforced (on z13)

zIIP SMT support – WLM considerations

- **Less overflow from zIIP to CPs** may occur because
  - zIIP capacity increases, and
  - number of zIIP CPUs double
- CPU time and CPU service **variability may increase**, because
  - Threads which are running on a core at the same time influence each other
  - Threads may be dispatched at thread density 1 or 2
    - Unlike other OS, z/OS attempts to dispatch threads densely
- Sysplex workload routing: Routing recommendation may change because of increased zIIP capacity and variability
- **Goals should be verified** for zIIP-intensive work, because
  - The number of zIIP CPUs double and the achieved velocity may change
  - "Chatty" (frequent dispatches) workloads may profit because there is a chance of more timely dispatching
  - More capacity is available
  - Any single thread will effectively run at a reduced speed and the achieved velocity will be lower. Affects processor speed bound work, such as single threaded Java batch

HiperDispatch "Unpark while capped"

- Previously, HiperDispatch

  - Parked all Vertical Low (VL) processors when a system capped via positive phantom weight
    - VLs are used for discretionary capacity and not required to absorb the LPAR weight
    - However, it was seen that, for some workloads, the reduced number of logical processors made it difficult to fully utilize the cap target capacity.
  - Unparked all VL processors when a system was capped by <u>negative phantom weight</u>, or some cases of PR/SM absolute capping

- Now, HiperDispatch can unpark VL processors <u>if</u> the processors can be used efficiently.

HiperDispatch refinement of "unused capacity" use

- HiperDispatch decisions consider the CPC-wide 'unused capacity share' situation

- The 'unused capacity share' calculation was enhanced to also include the LPAR configuration values
  - absolute capping value
  - negative phantom weight
  - number of logical processors
  - effective defined capacity and group capacity limit
  of possible 'unused capacity' receivers

CPC with 5 LPARs. LPAR1 has an absolute capping limit, which is indicated with the red line. LPAR2, and LPAR4 are unused capacity donors, while LPAR1 / 3 / 5 are unused capacity receivers.

Unused capacity

Guaranteed LPC capacity limit
Defined by LPAR current weight

Actual LPC capacity below the Guaranteed LPC capacity

Actual LPC capacity above the Guaranteed LPC capacity

White space
Unused LPC capacity of Guaranteed LPC capacity

- Absolute capping limit, or
- Negative phantom weight, or
- Limit caused by lack of LCPs

HiperDispatch refinement of "unused capacity" use

Enhanced unused capacity calculation

| LPAR 1 | LPAR 2 | LPAR 3 | LPAR 4 | LPAR 5 |

Today

Enhanced unused capacity calculation

- Figure on the left shows today's unused capacity calculation, which does not consider LPAR capping limits.

- Unused capacity calculation is only based on the receiver's weight share.

- Figure on the right shows an example of enhanced unused capacity calculation. It considers the capping limits of the receivers.

- Because LPAR1 is not able to use its total unused capacity share its 'not usable' unused capacity share portion increases the unused capacity share of LPAR5.

74

OA47968: HiperDispatch Optimizations for z13

- Vertical Low (VL) processors are used to absorb discretionary ("above the weight) processor capacity. VLs may float between different physical processors – consuming free physical capacity not used by other logical processors

- With OA47968 HiperDispatch takes benefit of the fact that lower VL numbers are likely to be topologically "closer" to the LPAR's VH and VM processors

- Visible effect is that the park time in the RMF CPU activity report should be increasing from the low to the high processor numbers
    - Due to weight changes numbers can still decrease

- On z13, even in the presence of free CPC capacity, unparking can be more restrictive, based on effective capacity used on the VM and VL processors.

Sample CPU Activity Report…
showing high VL numbers unparked

0---CPU---		---------------- TIME % ----------------				LOG PROC		--I/O INTERRUPTS--	
NUM	TYPE	ONLINE	LPAR BUSY	MVS BUSY	PARKED	SHARE %		RATE	% VIA TPI
0	CP	100.00	73.07	73.01	0.00	100.0	HIGH	331.0	47.48
...									
D	CP	100.00	62.53	62.49	0.00	100.0	HIGH	12768	14.71
E	CP	100.00	50.63	53.18	0.00	50.0	MED	134.8	60.51
F	CP	100.00	5.03	41.30	85.77	0.0	LOW	0.00	0.00
10	CP	100.00	5.14	38.64	84.88	0.0	LOW	0.00	0.00
11	CP	100.00	4.10	42.47	88.22	0.0	LOW	0.00	0.00
12	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
13	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
14	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
15	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
16	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
17	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
18	CP	100.00	8.81	46.39	76.66	0.0	LOW	0.00	0.00
19	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
1A	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
1B	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
1C	CP	100.00	0.00	-----	100.00	0.0	LOW	0.00	0.00
TOTAL/AVERAGE			33.37	62.25		1450		35779	15.49

Agenda

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview*
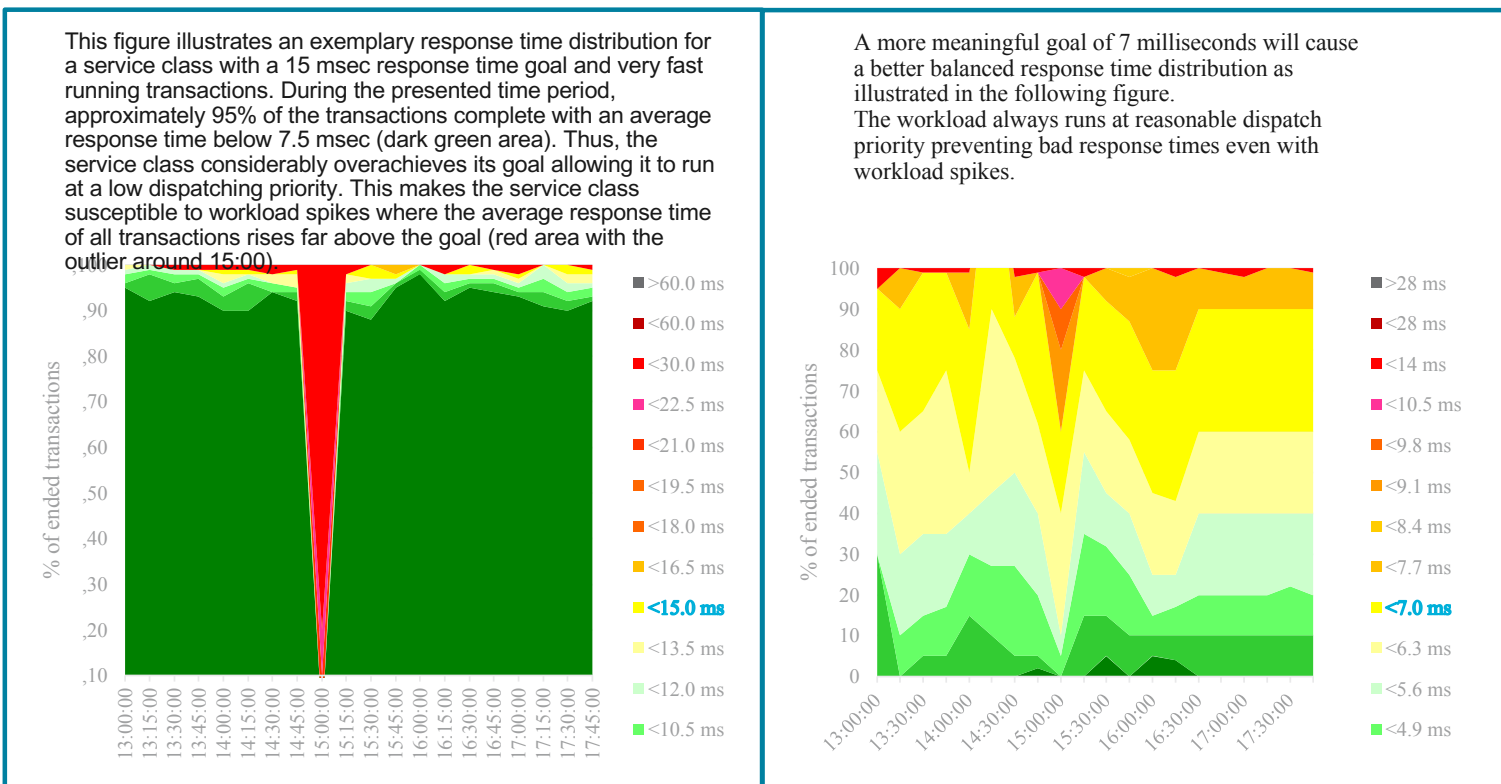
z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Shorter Response Time Goal*

- With z/OS V2R3, the current lower bound of 15 milliseconds for a response time goal is replaced by one millisecond allowing to specify meaningful goal values for very fast running transactions.

This figure illustrates an exemplary response time distribution for a service class with a 15 msec response time goal and very fast running transactions. During the presented time period, approximately 95% of the transactions complete with an average response time below 7.5 msec (dark green area). Thus, the service class considerably overachieves its goal allowing it to run at a low dispatching priority. This makes the service class susceptible to workload spikes where the average response time of all transactions rises far above the goal (red area with the outlier around 15:00).

A more meaningful goal of 7 milliseconds will cause a better balanced response time distribution as illustrated in the following figure.
The workload always runs at reasonable dispatch priority preventing bad response times even with workload spikes.

Legend (left chart): >60.0 ms, <60.0 ms, <30.0 ms, <22.5 ms, <21.0 ms, <19.5 ms, <18.0 ms, <16.5 ms, **<15.0 ms**, <13.5 ms, <12.0 ms, <10.5 ms

Legend (right chart): >28 ms, <28 ms, <14 ms, <10.5 ms, <9.8 ms, <9.1 ms, <8.4 ms, <7.7 ms, **<7.0 ms**, <6.3 ms, <5.6 ms, <4.9 ms

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Shorter Response Time Goal Usage & Invocation*

- The WLM Administrative Application is enhanced to allow the definition of service period response time goals below 15 milliseconds which is the minimum response time goal with z/OS V2R2 and below. The new minimum goal value is 0.001 seconds (one millisecond) and can be defined for base goals or when overriding attributes for a service class.

- When specifying an average response time goal, the total response time can be between 0.001 seconds (one millisecond) and 24 hours.

```
IWMAP34        Average response time goal

Enter a response time of up to 24 hours for period 2

Hours . . . . . 00  (0-24)
Minutes . . . . 00  (0-99)
Seconds . . . . 00.005  (0-9999)

Importance . . 2  (1=highest, 5=lowest)
Duration . . . 10000      (1-999,999,999, or
                           none for last period)


F1=Help     F2=Split    F5=KeysHelp  F9=Swap    F12=Cancel
```

```
IWMAP35        Response time with percentile goal

Enter a percentile and response time goal for period 3

Percentile . . 90  (1-99)

Hours . . . . . 00  (0-24)
Minutes . . . . 00  (0-99)
Seconds . . . . 00.008  (0-9999)

Importance . . 2  (1=highest, 5=lowest)
Duration . . . 10000      (1-999,999,999, or
                           none for last period)


F1=Help     F2=Split    F5=KeysHelp  F9=Swap    F12=Cancel
```

- With z/OS V2R3, the WLM Application level displayed on the Definition Menu panel changes to LEVEL035. As soon as at least one service class period is defined with a response time goal below 0.015 seconds (15 milliseconds), the Functionality level is raised to LEVEL035.

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

WLM Sysplex Routing Services Overview

WLM Interface	Typical Use (not exhaustive)	Purpose
IWMWSYSQ	Applications and sub-systems that want to consider free and displaceable capacity.	Obtain free & displaceable capacity of systems in Sysplex (1, 3, and 10 min rolling averages).
IWMSRSRS FUNCTION=SELECT FUNCTION=SPECIFIC	Sysplex Distributor BASEWLM DDF	Obtain best suited registered servers to route work to. Only capacity considered. Obtain list of registered eligible servers and recommended weights. Considers capacity goal achievement (PI), queue time for enclaves, health.
IWM4SRSC	Sysplex Distributor SERVERWLM	Obtain recommendation for a specific server address space. No registration required. Capacity, server-specific capacity goal achievement (PI), abnormal termination rate, health is considered; optionally crossover cost and importance level weighting

Routing Enhancements for Softcapping*

- ➢ WLM Sysplex Routing services like IWMSRSRS and IWM4SRSC base their recommendation on the free and displaceable capacity of the systems in the Sysplex (3 minutes rolling average of actual capacities).
- ➢ This might result in routing work to a system which will be capped shortly thereafter due to Defined Capacity Limit or Group Capacity Limit.
- ➢ WLM is enhanced to **take the capping limits into account** when the free/displaceable capacity is determined. WLM will calculate the estimated time to capping for a system. The closer the system is to capping the more the available capacity will be reduced by the specified limit and influence the routing recommendations to send less work to the system.

New IEAOPTxx parmlib option RTCapLeadTime*

IEAOPTxx RTCAPLEADTIME=nn	
<u>0</u>	The default behavior is as today: capping of this system will not be considered in advance.
[1-60]	Specifies the time in minutes, how long in advance an upcoming soft capping should influence WLM's Sysplex routing recommendations. When the estimated time to capping is less than n minutes WLM will consider the upcoming soft capping in its routing recommendations.

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Agenda

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview

z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Dependent Job Control for JES2

- JES2 in z/OS V2.2 provides a new job scheduling scheme similar to "JES3' Dependent Job Control" which in turn allows for a set of concurrent jobs to be run

```
//JGROUP   JOBGROUP
//INITJ    GJOB
//CONC1    GJOB
//         AFTER NAME=INITJ
//CONC2    GJOB
//CONC3    GJOB
//         CONCURRENT NAME=(CONC1,CONC2)
//ENDJ     GJOB
//         AFTER NAME=CONC1,WHEN=(RC=0)
//JGROUP   ENDGROUP
//INITJ    JOB TIME=NOLIMIT,REGION=0K,MSGCLASS=H,CLASS=J
//         SCHEDULE JOBGROUP=JGROUP
//STEP1    EXEC  PGM=...
//CONC1    JOB TIME=NOLIMIT,REGION=0K,MSGCLASS=H,CLASS=J
//         SCHEDULE JOBGROUP=JGROUP
//STEP1    EXEC  PGM=...
...
```
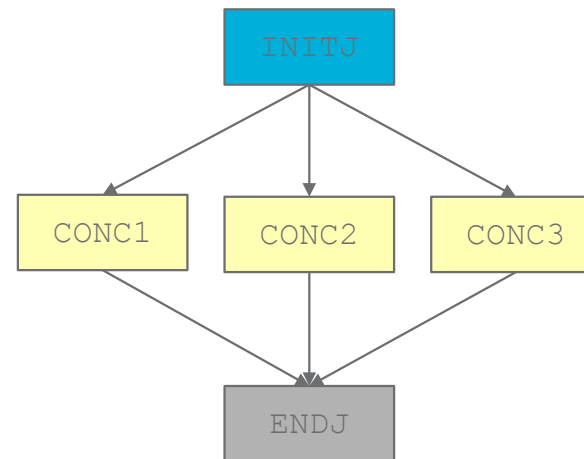
Dependent Job Control for JES2

- JES2 in z/OS V2.2 provides a new job scheduling scheme similar to "JES3' Dependent Job Control" which in turn allows for a *set of concurrent jobs* to be run

- For a set of concurrent jobs, WLM extends the *demand batch initiator* interface with JES2:

  - WLM returns the most eligible system for starting the demand batch initiators, or indicates that all candidate systems are too constrained

  - If a system is eligible, then
    - WLM reuses drained initiators, or
    - starts demand batch initiators.
    Both select the concurrent jobs specified by JES2

  - When the jobs are finished, both the reused and the newly started initiators go to the drained state

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

IWM4OPTQ:
Interface to retrieve IEAOPT keywords and values

- Prior to this support it was difficult for a monitoring product to identify
  - what IEAOPTxx parameters are known to the system, and
  - the values for those parameters
- IWM4OPTQ returns a self-describing list of all keywords, and their values. Return area is mapped by IWMWOPTI
- Exploited e.g. by RMF

```
>>-+------+--IWM4OPTQ-------OPTINFO_BLOCK=optinfo_block--------->
   '-name-'


>---,ANSLEN=anslen----------QUERYLEN=querylen------------------->


>--+----------------+--+----------------+----------------->
   '-,RETCODE=retcode-'  '-,RSNCODE=rsncode-'

   .-,PLISTVER=IMPLIED_VERSION-.
>--+--------------------------+--------------------------->
   +-,PLISTVER=MAX------------+
   '-,PLISTVER=0--------------'


   .-,MF=S-------------------------.
>--+-------------------------------+-------------------><
   |                  .-,0D---.    |
   +-,MF=(L-,list addr-+-------+-)-----+
   |                  '-,attr-'   |
   |                  .-,COMPLETE-.|
   '-,MF=(E-,list addr-+----------+-)-'
```

IWM4HLTH:
Extensions for health based routing

- WLM Sysplex routing services provide advice for routing work within a Sysplex
  - Enable distributed client/server environments to balance work among multiple servers based, on capacity, performance, server health
  - Utilized e.g. by Sysplex distributor (SERVERWLM), DB2 DDF

- The IWM4HLTH service allows to modify the health value when the health status of the server changes for the worse or better

- Before z/OS V2.2 the server health value solely is based on self-assessment with only the last value reported is being kept by WLM

- **With z/OS V2.2 the IWM4HLTH service is extended to work with multiple components providing their views of the health of a server address space.**
  The new IWM4QLTH service allows to query the health.

IWM4HLTH: Extensions for health based routing

```
                                              +--,FUNCTION=SET----+
                                              |                   |
>>-+------+---IWM4HLTH---STOKEN=stoken--+-------------------+--,HEALTH=health-------->
   |      |                             |                   |
   +-name-+                             +--,FUNCTION=RESET--+


  +--,SUBSYS=NO_SUBSYS--+   +--,SUBSYSNM=NO_SUBSYSNM--+   +--,HEALTHRSN=NO_RSN-----+
  |                     |   |                         |   |                        |
>-+---------------------+---+-------------------------+---+------------------------+-->
  |                     |   |                         |   |                        |
  +--,SUBSYS=subsys-----+   +--,SUBSYSNM=subsysnm-----+   +--,HEALTHRSN=healthrsn--+


>-+-------------------+---+-------------------+------------------------------------>
  |                   |   |                   |
  +--,RETCODE=retcode--+   +--,RSNCODE=rsncode--+


  +--,PLISTVER=IMPLIED_VERSION--+   +--,MF=S-----------------------------+
  |                             |   |                                    |
>-+-----------------------------+---+------------------------------------+--------><
  |                             |   |                                    |
  +--,PLISTVER=MAX--------------+   |                       +-,0D---+     |
  |                             |   +--,MF=(L-,list addr-+-------+-)------+
  +--,PLISTVER=0----------------+   |                     +-,attr-+       |
  |                             |   |                                     |
  +--,PLISTVER=1----------------+   |                       +-,COMPLETE-+ |
                                    +--,MF=(E-,list addr-+-----------+-)--+
```

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

Runtime Diagnostics of Health Values

- z/OS Runtime Diagnostic display & analyze all server health values <100 through F HZR,ANALYZE command:

```
-----------------------------------------------------------------
EVENT 17: HIGH - SERVERHEALTH - SYSTEM: SYS1     2016/04/19 - 08:00:30
JOB NAME: DB1XDIST  ASID: 01CC  CURRENT HEALTH VALUE: 75
CURRENT LOWEST HEALTH VALUES:
            SUBSYSTEM    HEALTH                    REPORTED
 SUBSYSTEM  NAME         SETTING         REASON   DATE AND TIME
 DB1TDIST                75                        2016/04/19 06:01:04
  ERROR: ADDRESS SPACE SERVER CURRENT HEALTH VALUE LESS THAN 100.
  ERROR: THIS VALUE MAY IMPACT YOUR SYSTEM OR SYSPLEX TRANSACTION
  ERROR: PROCESSING.
 ACTION: USE YOUR SOFTWARE MONITORS TO INVESTIGATE THE ASID AND TO
 ACTION: DETERMINE THE IMPACT OF THE HEALTH OF THE ADDRESS SPACE TO
 ACTION: OVERALL TRANSACTION PROCESSING.
-----------------------------------------------------------------
```

WLM-managed DB2 Bufferpools: Overall flow

- DB2 registers bufferpool with WLM

- WLM will recommend to grow the size of the bufferpool when the Performance Index of a Service Class Period is impacted and bufferpool delays are a significant contributor

- WLM will recommend to shrink the size of the bufferpool due to donation to a suffering Service Class period, or due to regular housekeeping cycles

- DB2 de-registers bufferpool from WLM management
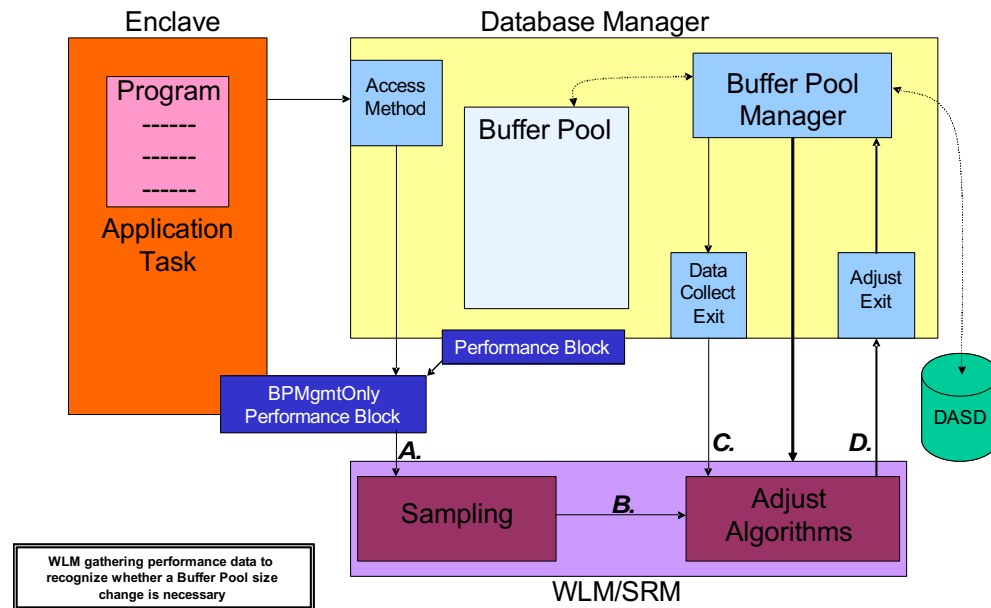
- DB2 10: ALTER BUFFERPOOL [VPSIZE(s)] AUTOSIZE(YES)
  → MIN size = 0.75 x VPSIZE
  → MAX size = 1.25 x VPSIZE

- DB2 11: MIN/MAX can be specified

- Initial USED size between MIN size and MAX size

- Management range between MIN and MAX sizes

WLM-Managed DB2 Bufferpool: Changes in z/OS V2.2 plus V2.1

A bufferpool can be increased when

- Performance index impacted and buffer pool delays are a significant contributor

A bufferpool may shrink…

- Due to donation to a suffering service class period
  - May suffer storage related delays

- Due to regular housekeeping cycles
  - Consider one BP reduction candidate per 10 sec interval
    - BP idle - had no references
    - No delays, i.e. 100% hit ratio
    - Least important period showing buffer pool delays
    Any bufferpool may shrink no more than once per 5 min

- When WLM recommends to increase the size of a bufferpool, DB2 accepts the recommended size as new current VPSIZE. DB2 does not necessarily use the entire recommended size ➔ Used size of the bufferpool will be less or equal the current VPSIZE
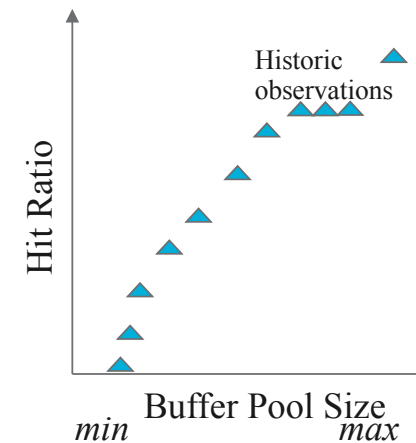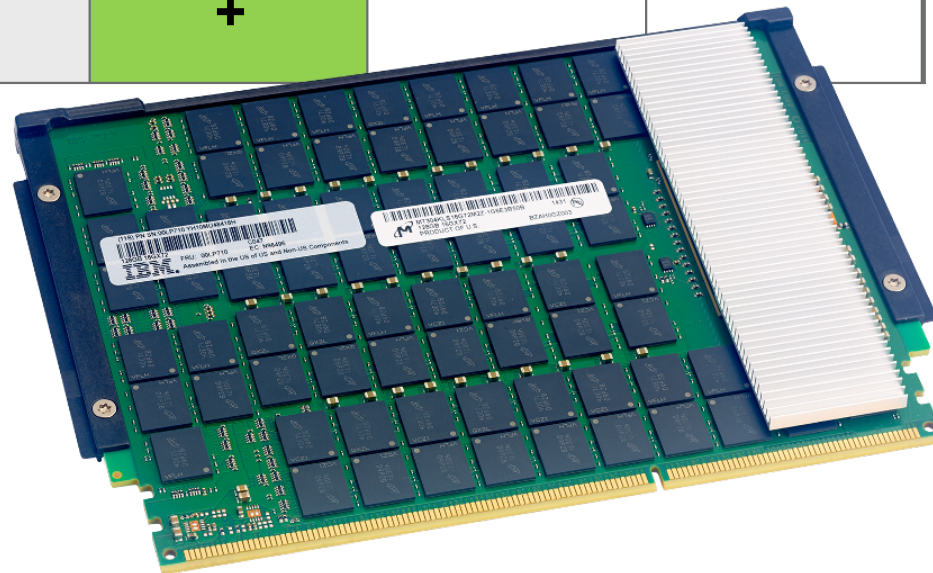
Bufferpool
BP: BP11   ,0100002000000002,AS: DBX2DBM1,SC: $SRMS059

CurSize — TarSize — MinSize — MaxSize — UseSize

95

© Copyright IBM Corp. 2017
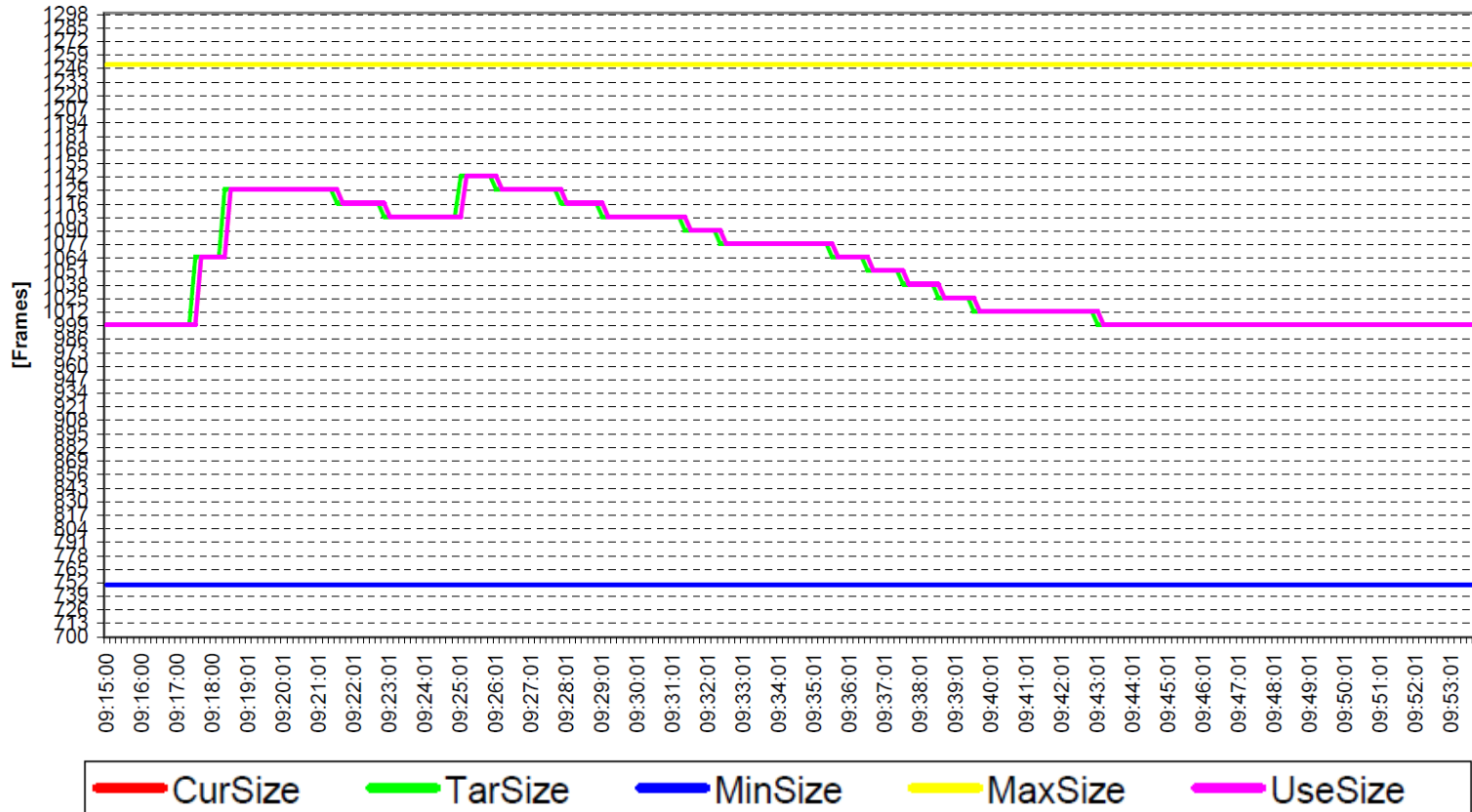
SRM Enhancements for large real storage

z/OS release Function	V2.2	V2.1	V1.13
MCCFXPTR Limit	+	OA44668 OA44207(RSM)	
New AUTO keyword for RCCFXET and RCCFXTT	+		

New MCCFXPTR Limit

- The MCCFXTPR keyword in the IEAOPTxx specifies the percentage of online storage that may be page fixed before a pageable storage shortage is detected and message IRA400E is issued.

  - Before OA44668, MCCFXTPR default of 80% requires that 20% (100 minus MCCFXTPR) of storage remain pageable, regardless of the amount of online storage. On systems with large amounts of central storage, the MCCFXTPR default of 80% can result in a pageable storage shortage being detected when there is still plenty of pageable storage.

  - With OA44688 **at most 64GB** of pageable online storage will be required before a pageable storage shortage is recognized.

800G   1TB
80%
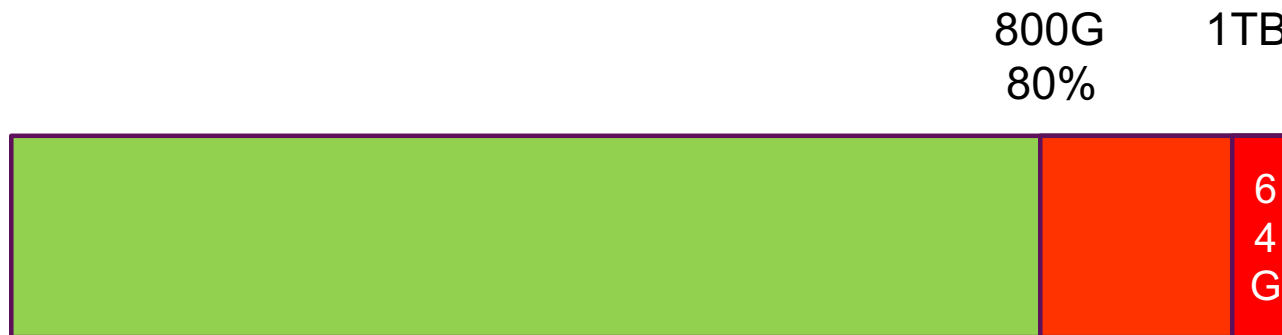
New AUTO keyword for RCCFXETand RCCFXTT

- The IEAOPTxx *RCCFXTT* keyword specifies low and high threshold of fixed real storage:
  - SRM uses these thresholds to determine if the **system MPL** needs to be increased/decreased. The default is 66% and 72%.
  - On small systems such percentages are not a problem.
  - On a 1TB LPAR these percentages imply that WLM will stop increasing the MPL. when 660G of storage is fixed

- Similarly, *RCCFXET* specifies the low and high thresholds of fixed real storage below 16M. SRM uses these thresholds to determine if the **system MPL** needs to be increased/decreased. The default is 82% and 88%.
  - This OPT keyword is also enhanced, mainly for consistency with the RCCFXTT keyword. The default is still: RCCFXET=(82, 88)

- Both keywords were enhanced to accept a value of **AUTO**
  - AUTO allows SRM to compute thresholds based on available storage.
  - Allows to higher utilize available storage in large systems without risking system shortages
  - AUTO needs to be specified in IEAOPTxx (not default)

New AUTO keyword for RCCFXET and RCCFXTT

Current RCCFXTT thresholds:

660G 66% 720G 72% 1TB
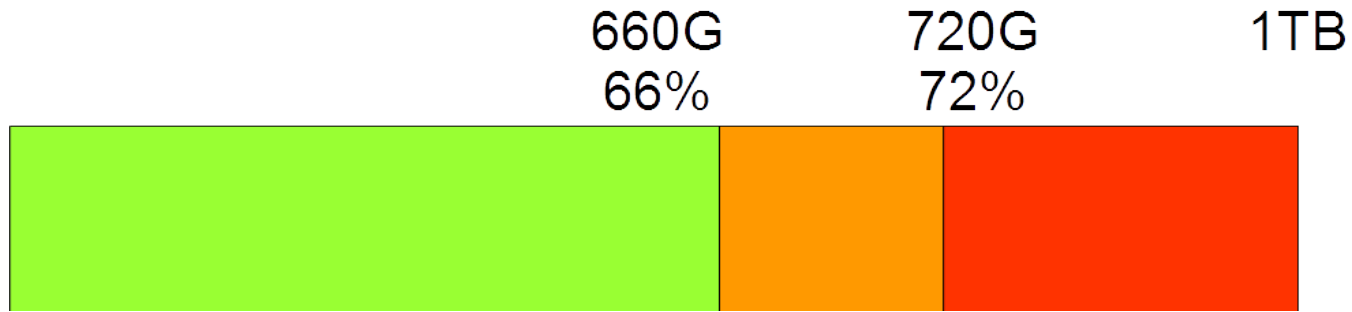
Current RCCFXET thresholds:
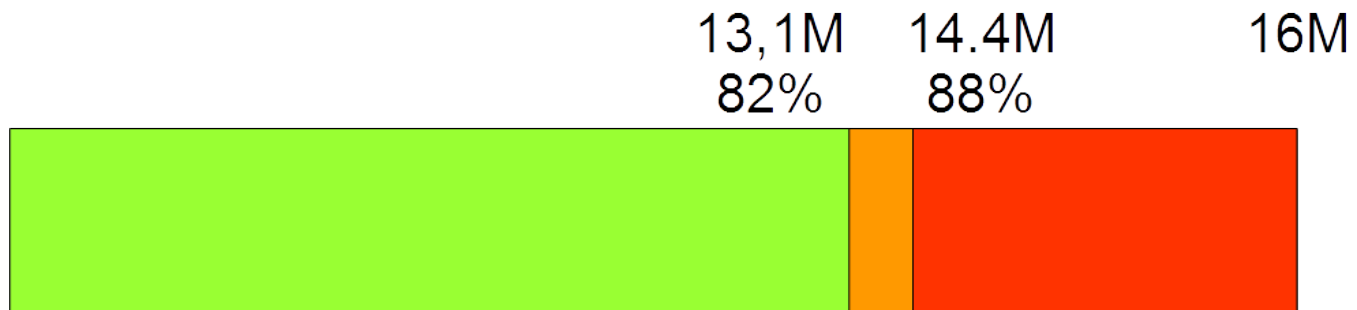
13,1M 82% 14.4M 88% 16M

Agenda

New capping options

Reporting enhancements for CICS and IMS, and Mobile Workloads

Container Pricing*

More granular resource controls*

IBM z14 Exploitation

IBM z13 Support

z/OS V2.3 Preview

z/OS V2.2 Enhancements

Other service stream enhancements and recommendations

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

OA50900 – WLM Absolute MSU Capping requires zEC12 GA2 or later

- The initial documentation for Absolute MSU capping did not indicate that a zEC12 GA2 or later processor is required to enable this function.

- DOC APAR OA50900 is going to document this requirement:

IEAOPTxx ABSMSUCAPPING=	
NO	Defined capacity limits and group capacity limits should be enforced only while the long term four hour rolling average consumption exceeds the respective limit (existing and usually desired behavior).
YES	Defined capacity limits and group capacity limit should be enforced **permanently, independently of the long term four hour rolling average consumption**. Becomes effective on zEC12 GA2 or later processors.

OA50403 – Additional Message IWM068I when running at reduced processor speed

- When the effective processor speed changes WLM issues
  `IWM063I WLM POLICY WAS REFRESHED DUE TO A PROCESSOR SPEED CHANGE OR MT MODE CHANGE`, and identifies the specific reason via IWM064I. For example, `IWM064I THE SYSTEM IS RUNNING WITH REDUCED CAPACITY BECAUSE OF A MANUAL CONTROL SETTING` (I.e., static power save mode)

- IWM064I is also issued when a system is IPL'ed while running at reduced capacity

- Problem: Automation would not yet be available during early IPL. Alerts may be missed.

- APAR OA50403 (z/OS V2.2) provides new message IWM068I that is issued
  – hourly after IWM064I,
  – approximately 15 min after IPL

- For example:
  `IWM068I THE SYSTEM IS STILL RUNNING WITH REDUCED CAPACITY BECAUSE OF A MANUAL CONTROL SETTING`
  Other possible reasons:
  – `THE SYSTEM IS STILL RUNNING WITH REDUCED CAPACITY BECAUSE OF A MACHINE EXCEPTION CONDITION`
  – `THE SYSTEM IS STILL RUNNING WITH REDUCED CAPACITY BECAUSE OF A NON-EXCEPTION MACHINE CONDITION`
  – `THE SYSTEM IS STILL RUNNING WITH REDUCED CAPACITY BECAUSE OF AN EXCEPTION CONDITION EXTERNAL TO THE MACHINE.`

105

OA50291 – New option to eliminate ICH70001I messages for WLM Application Environment Servers

- With Security=User, use of WLM application environments can result in a huge number of ICH70001I xxxxxx LAST ACCESS AT 10:38:32 ON THURSDAY, NOVEMBER 15, 2015 messages
  - Fill up logs
  - RACF data base updates

- WLM APAR OA50291 allows for suppressing the RACF message and reducing the RACF data base updates

- To enable daily logon statistics it is required to
  - Update security (RACF) definitions
  - Specify new IEAOPTxx keyword: SuppSAFinfoMsg=YES

IEAOPTxx SuppSAFinfoMsg =	
NO	WLM performs a standard verification of the application environment server address spaces through the security product.
YES	WLM performs a verification of the application environment server address spaces through the security product and passes an APPL class profile with a name equal to the procedure name associated with the application environment. SAF informational messages, such as RACF message ICH70001I will be suppressed.

107

OA50291 – Option to eliminate ICH70001I messages
- Sample security definitions-

- The following security definitions allow to reduce the impact of recording logon statistics by recording for only the first daily logon by each user, rather than for every logon by each user:
  - Define a RACF APPL profile
    RDEFINE APPL *applname* UACC(NONE) APPLDATA('RACF-INITSTATS(DAILY)')
    where *applname* is the name of the JCL procedure name used for the WLM application environment
    - Ensure that no such APPL already exists
    - Generic profiles may be used
  - If UACC(NONE) was specified on the APPL, permit users READ access to the APPL profile
    - PERMIT *applname* CLASS(APPL) ID(*userid*) ACCESS(READ)
  - Activate or refresh the APPL profile:
    - if not yet active, issue
      SETROPTS CLASSACT(APPL) RACLIST(APPL)
    - SETROPTS RACLIST(APPL) REFRESH

- For more details refer to Reducing application logon statistics (https://ibm.biz/Bd4S3s)

OA48161 – Service Stream Enhancements for reduced WLM Address Space utilization

Minimizing Sampling Overhead for Performance Block (PBs)

- In some environments a – relatively – higher percentage of WLM address space utilization may be observed

- Typically these are mostly idle test systems hosting many CICS or DB2 subsystems
  - In such cases the WLM task responsible for sampling Performance Blocks (PBs) allocated by theses subsystems may incur the highest CPU cost

- Recommendations:
  - On z/OS releases up to z/OS V1.13 install PTF for OA38280
    - Eliminates an IVSK instruction that is more costly on IBM z196 and later hardware
  - On z/OS V2.1 install PTF for OA48161
    - Eliminates an MVCSK instruction that was introduced with z/OS V2.1
  - Reduce the number of PB control blocks allocated by the subsystems:
    - CICS: Max task (MXT)
    - DB2 MSTR+DBM1:   1000 + MAXDBAT
                                          + (n · 500) as needed
    - Reducing the effective max task and MAXDBAT numbers can help to reduce the WLM address space consumption

Service Stream Enhancements for more aggressive Blocked Workload support (OA44526)

- Problem addressed:
  - The current minimum value that can be specified for the Blocked Workload interval threshold BLWLINTHD is 5 sec.
  DB2 could profit from earlier or more frequent trickling.

- More aggressive specifications will be enabled by OA44526

  - New lower limit is 1 sec

  - BLWLINTHD default and BLWLTRPCT remain unchanged
    - Consider lowering BLWLTRPCT with very small BLWLINTHD values if amount of trickle cycles that may be handed out is a concern.

XML Format WLM service definitions recommended

▪ For several releases WLM has supported service definitions in **XML format**
  – z/OSMF WLM task
  – ISPF Administrative Application: "Save as XML"…

▪ XML format avoids particular problems with the ISPF tables format, namely coexistence behavior, when a new functionality level needs to be introduced, and the number of table columns needs to be extended.

▪ **Recommendation**:
  Consider using the XML-format for your WLM service definition data sets.

z/OS Workload Management
- More Information -

- z/OS WLM Homepage:

  http://www.ibm.com/systems/z/os/zos/features/wlm/
  - Inside WLM: https://ibm.biz/BdF4L4

- z/OS MVS documentation
  - z/OS MVS Planning: Workload Management:
    http://publibz.boulder.ibm.com/epubs/pdf/iea3w101.pdf
  - z/OS MVS Programming: Workload Management Services:
    http://publibz.boulder.ibm.com/epubs/pdf/iea3w201.pdf

- *IBM Redbooks publication*s:
  - System Programmer's Guide to: Workload Manager:
    http://publib-b.boulder.ibm.com/abstracts/sg246472.html?Open
  - ABCs of z/OS System Programming Volume 12
    http://publib-b.boulder.ibm.com/abstracts/sg247621.html?Open
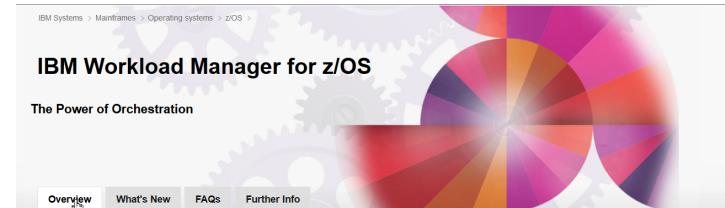
WLM Topology Report Tool (As-is)

- New *as-is* tool available for download from the WLM homepage
  - https://ibm.biz/BdE74v

- Visualizes mapping of HiperDispatch affinity nodes to physical structure

- Supports IBM zEC10 and later

- To use:
  1. Download from above location
  2. Run installer
  3. Upload Host code to a z/OS system
  4. Collect SMF99.14 records

Sample output (zEC12):

Topology for 07-21-2014-13:44:27 ,System: IRD9

Book_2						
Chip_1 IRD9_01_MCPU002 IRD9_01_MCPU003 IRD9_02_MAAP008 IRD9_03_MIIP009 IRD9_01_MCPU001 IRD9_01_LCPU002 IRD9_02_MAAP008 IRD9_03_MIIP009		**Chip_3** IRD9_01_LCPU004 IRD9_01_LCPU005 IRD9_01_LCPU006	**Chip_4** IRD9_01_LCPU007 IRD9_01_LCPU010 IRD9_01_LCPU011 IRD9_01_LCPU003 IRD9_01_LCPU004 IRD9_01_LCPU005	**Chip_5** IRD9_01_LCPU006 IRD9_01_LCPU007 IRD9_01_LCPU010	**Chip_6** IRD9_01_HCPU000 IRD9_01_HCPU001 IRD9_01_HCPU000 IRD9_01_LCPU011