RMF – The Latest and Greatest

Peter Muench (pmuench@de.ibm.com)
IBM Corporation

Thursday, March 5, 2015
Session 16816

#SHAREorg

SHARE is an independent volunteer-run information technology association that provides education, professional networking and industry influence.

Permission is granted to SHARE Inc. to publish this presentation paper in the SHARE Inc. proceedings; IBM retains the right to distribute copies of this presentation to whomever it chooses.

© Copyright IBM Corp. 2015

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®

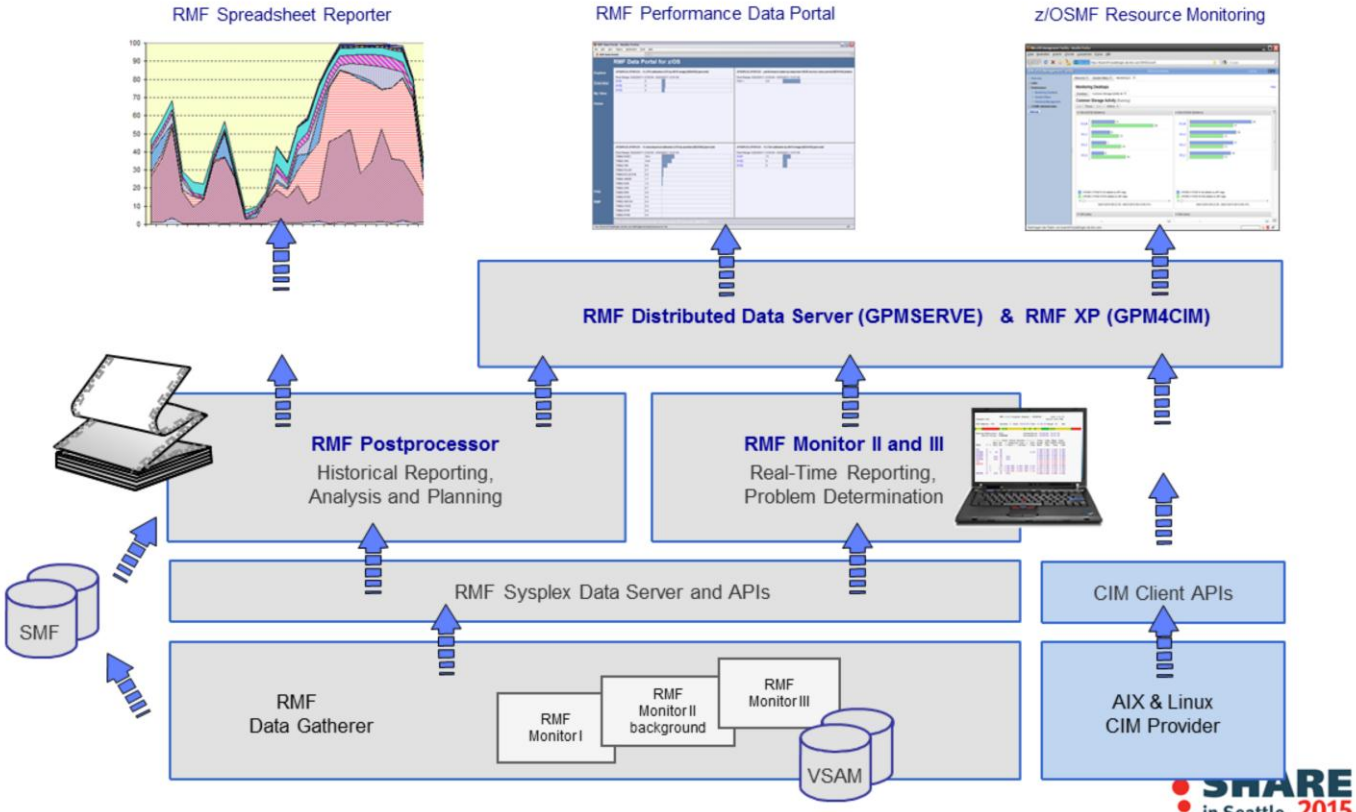
The following are trademarks or registered trademarks of other companies.

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.

RMF Product Overview



- z/OS Resource Measurement Facility (RMF) is an optional priced feature of z/OS. It supports installations in performance analysis, capacity planning, and problem determination. For these disciplines, different kinds of data collectors are needed:
  - Monitor I long term data collector for all types of resources and workloads. The SMF data collected by Monitor I is mostly used for capacity planning and performance analysis
  - Monitor II snap shot data collector for address space states and resource usage. A subset of Monitor II data is also displayed by the IBM SDSF product
  - Monitor III short-term data collector for problem determination, workflow delay monitoring and goal attainment supervision. This data is also used by the RMF PM Java Client and the RMF Monitor III Data Portal
- Data collected by all three gatherers can be saved persistently for later reporting (SMF records or Monitor III VSAM datasets)
- While Monitor II and Monitor III are realtime reporters, the RMF Postprocessor is the historical reporting function for Monitor I data
- One of the key components for the sysplex wide access of Monitor III data is the RMF Distributed Data Server (DDS). Beginning with RMF for z/OS 1.12, DDS supports HTTP requests to retrieve RMF Postprocessor data from a selection of RMF Postprocessor reports. Since the requested data are returned as XML document, a web browser can act as Data Portal to RMF Postprocessor data.
- Since z/OS 1.12 there's another exploiter of the RMF DDS data: The z/OSMF Resource Monitoring plugin of the z/OS Management Facility.
- RMF for z/OS 1.13 enhances the DDS layer with a new component:
  - RMF XP is the new solution for Cross Platform Performance Monitoring
  - Provides a seamless performance monitoring for all operating systems running on the IBM zEnterprise Bladecenter Extension.

RMF Enhancements at a Glance

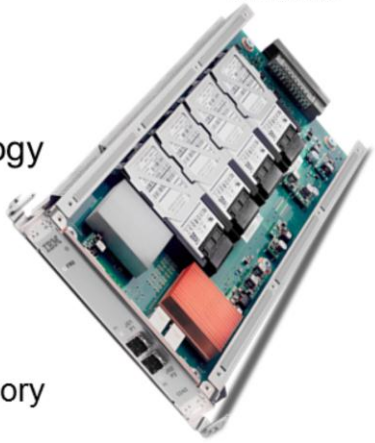
- IBM zEnterprise EC12 Support
  - Statistics for Flash Memory and Pageable Large Pages
  - Support of Coupling Facility Flash Memory
  - z Enterprise Data Compression Express Reporting (zEDC)
  - Shared Memory Communication Reporting (SMC-R)
  - Support of I/O Interrupt Delay Time Facility
  - Extended Infiniband Link Reporting
  - Support of Crypto Express4 Card
- IBM z13 Support
  - RMF support for SMT, Crypto Express5S …
  - PCIe reporting enhancements
- Exploitation of System z Integrated Information Processors
  - zIIP Usage Option for Monitor III Gatherer
- Postprocessor XML formatted Reports
  - Transition to XML Format almost complete
  - Advanced Sorting and Filtering Capabilities
- RMF XP Enhancements
  - SMF Type 104 Recording for all Platforms
- z/OSMF Resource Monitoring
  - Context sensitive Application Linkage to WLM
  - Historical Reporting & Spreadsheet Export

- In accordance with the availability of new z/OS releases and new hardware functionality, the capabilities of RMF
- are enhanced consecutively
- Together with the new zEnterprise EC12 servers RMF provides the following support:
  - Usage statistics for Pageable Large Pages and Storage Class Memory in various RMF Postprocessor, Monitor II and Monitor III reports.
  - RMF Support for CFCC Storage Class Memory exploitation.
  - New statistics to report PCIE based data compression and shared memory communication.
  - Reporting of new I/O Interrupt Delay Time
  - Additional channel path information for channel paths of type CIB or CFP
  - Support of Crypto Express4 Card and the Warning-Track-Interruption-Facility
- Various RMF new function APARs exploit the functionality of the IBM z13.
- With z/OS V2R1 RMF, the Monitor III Data Gatherer (RMFGAT) can partially offload work to zIIP processors
- Starting with z/OS V1R11, a subset of RMF Postprocessor reports can be generated in XML format. The new format provides:
  - Browser based, state-of-the-art display of SMF Type 7x data
  - Standardized format for access to performance data via XML parsing
- RMF for z/OS V2R1 completed the work on enabling Postprocessor reports for XML format output and provides sorting and filtering capabilities when the XML report is displayed in a web browser.
- RMF XP is the new solution for Cross Platform Performance Monitoring
  - Provides a seamless performance monitoring for all operating systems running on the IBM zEnterprise Bladecenter Extension.
  - With z/OS V2R1RMF XP can store the cross platform performance data in the new SMF record type 104.
- zIOSMF Resource Monitoring enhancements:
  - Context sensitive Application Linkage to z/OSMF WLM plugin
  - Display historical data and export the data as CSV for spreadsheet usage.

zEC12 – Flash Express

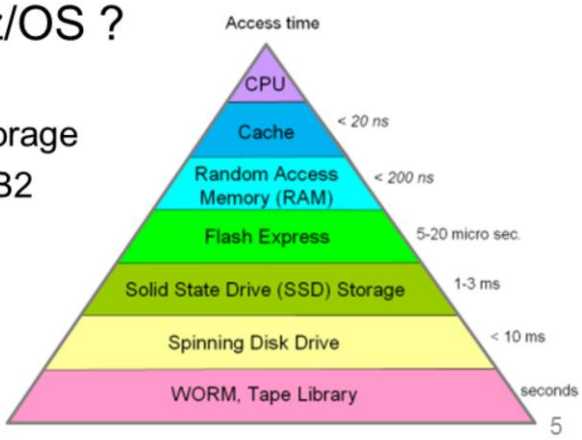
What is Flash Express?

- New memory hierarchie of the zSeries family
- Delivers tier within the fast Solid State Drive (SSD) technology
- Also denoted as Storage Class Memory (SCM)
- Integrated on PCI Express attached RAID 10 Cards
  - Packaged as two card pair
  - Each card holds 1.4 TB of memory per mirrored card pair
  - Maximum value of four card pairs delivers up to 5.6 TB of memory



How is Flash Express exploited by z/OS ?

- Designed for improved paging performance
- Compelling addition to traditional auxilliary storage
- Supports Pageable Large Pages, e.g. with DB2 and Java workloads
- Eliminates delays from SVC or standalone dump processing
- Provides overflow capability for MQSeries list structures



- The Flash Express feature (FC 0402) of IBM zEnterprise EC12 (zEC12) server is a new memory hierarchie of the zSeries family
- It consists of non-volatile storage using solid state devices on a PCIe card form factor.
- Flash Express implements a new tier of memory, called Storage Class Memory (SCM).
- Flash Express Cards are installed in pairs, which provides mirrored data to ensure a high level of availability and redundancy.
- In each Flash Express card, the data is stored in four solid-state disks in a RAID configuration. If a solid-state disk fails, the data are reconstructed dynamically. The cards in a pair mirror each other over a pair of cables, in a RAID 10 configuration. If either card fails, the data is available on the other card.
- Each Flash Express card has a capacity of 1.4 TB of usable storage.
- A maximum of four pairs of cards can be installed on a zEC12, for a maximum capacity of 5.6 TB of storage.
- Flash Memory is assigned to partitions like Main Memory from the allocation panel on the zEC12 Service Element (SE)
- z/OS can use Flash Express storage as Storage Class Memory (SCM) for paging.
- Flash Express helps to improve paging performance since page access time from Flash Express is faster than from DASD devices.
- The z/OS paging subsystem can work with a mix of Flash Express storage and External Disk.
- z/OS detects whether Flash Express storage is assigned to the LPAR and will try to page to Flash Express before using paging datasets on DASD.
- In combination with the new pagebale 1MB pages, Flash Express helps to improve the performance of DB2 and Java workloads.
- Latency delays in SVC or standalone dump processing caused by page-ins from DASD can be significantly reduced by Flash Express.
- The Coupling Facility can exploit SCM as overflow capacity for list structure data. This functionality can be used by MQSeries to avoid structure-full conditions.

zEC12 – Flash Memory & Pagegable Large Pages RMF Support

OA38660

- New Storage Class Memory (SCM) statistics in
  - RMF Postprocessor Paging Activity report
  - RMF Postprocessor Page Data Set Activity (PAGESP) report
  - RMF Monitor II Page Data Set Activity (PGSP) report
- New statistics for Pageable Large Pages in
  - RMF Postprocessor Paging Activity report
  - RMF Postprocessor Virtual Storage Activity (VSTOR) report
  - RMF Monitor III Storage Memory Objects (STORM) report

- RSM web deliverable (FMID JBB778H) exploits the new flash memory by introducing a new tier of memory called Storage Class Memory (SCM) that is used by the paging subsystem.
- In addition to support for the existing large (1 MB) pages and frames, zEC12 supports pageable large pages when SCM is configured and allocated to z/OS.
- With RMF new function APAR OA38660, RMF provides the SCM and Pagegable Large Pages support for z/OS 1.13
- The support enhances RMF Postprocessor, Monitor II and Monitor III reports with various new statistics for SCM and pageable large pages.

zEC12 – Flash Memory & Pagegable Large Pages…

⇨ New SCM statistics in the FRAMES AND SLOT COUNTS section of the RMF Postprocessor Paging Activity report

The number of shared pages backed on SCM

```
SHARED FRAMES            TOTAL SLOTS       CENTRAL STORAGE  FIXED TOT  FIXED BEL   AUX DASD    AUX SCM
----------------------   -----------  --------------------  ---------  ---------  ---------  ---------
      MIN                      7,937                    44         30          0         13          0
      MAX                      7,937                    44         30          0         13          0
      AVG                      7,937                    44         30          0         13          0

LOCAL PAGE DATA SET SLOTS      TOTAL   AVAILABLE        BAD    NON-VIO        VIO
-------------------------  ---------  ---------  ---------  ---------  ---------
      MIN                  5,399,997  4,269,302          0  1,128,251          0
      MAX                  5,399,997  4,271,746          0  1,130,695          0
      AVG                  5,399,997  4,269,838          0  1,130,159          0

SCM PAGING BLOCKS              TOTAL   AVAILABLE        BAD     IN-USE
-------------------------  ---------  ---------  ---------  ---------
      MIN                          0          0          0          0
      MAX                          0          0          0          0
      AVG                          0          0          0          0
```

System wide statistics of 4K SCM paging blocks as: Total, Available, Unavailable and Used 4K blocks

- The Frame and Slot Counts section of the RMF Postprocessor Paging Activity report is enhanced to report a new SCM Paging Block with statistics about the system wide usage of SCM.
- If SCM is not available the SCM Paging Block section is not shown.
- The SHARED FRAMES statistics are extended with a new metric to display the number of shared pages that are backed on SCM.

zEC12 – Flash Memory & Pagegable Large Pages...

- New SCM and Large Pages statistics in the MEMORY OBJECTS section of the RMF Postprocessor Paging Activity report

```
                                    P A G I N G   A C T I V I T Y

OPT = IEAOPT00   LFAREA SIZE = 209715200     MEMORY OBJECTS AND HIGH VIRTUAL STORAGE FRAMES
-------------------------------------------------------------------------------------------------------------------------

MEMORY OBJECTS          COMMON     SHARED       1 MB
------------------   ---------  ---------  ---------
      MIN                   53          1          1
      MAX                   56          1          4
      AVG                   54          1          2

1 MB FRAMES          ------------ FIXED ------------  ----------- PAGEABLE ----------
------------------       TOTAL  AVAILABLE     IN-USE      TOTAL  AVAILABLE     IN-USE
      MIN                  200         80         30        560        496         57
      MAX                  200        170        120        560        503         64
      AVG                  200        136         64        560        501         59

HIGH SHARED FRAMES       TOTAL      CENTRAL STORAGE            AUX DASD    AUX SCM
------------------   ---------      ---------------          ---------  ---------
      MIN            136902.1M                 206                   0          0
      MAX            136902.1M                 206                   0          0
      AVG            136902.1M                 206                   0          0

HIGH COMMON FRAMES       TOTAL                                  AUX DASD    AUX SCM
------------------    --------                                  ---------  ---------
      MIN             17301504                                        0          0
      MAX             17301504                                        0          0
      AVG             17301504                                        0          0
```

System wide usage of Large Frame Area (Fixed Frames) and Pageable Large Frames

Size of high virtual **shared** and **common** area in units of 4KB pages

Number of auxiliary storage slots used for high virtual common and shared memory pages that are backed on SCM storage

- The Memory Objects and High Virtual Storage Frames section of the Postprocessor Paging Activity report now shows values for both fixed and pageable 1 MB frames
- New statistics for High Shared and High Common Frames:
  - TOTAL: Size of high virtual shared and common area in units of 4 KB pages.
  - AUX SCM: Number of auxiliary storage slots used for high virtual shared and common pages that are backed on SCM storage.
- The number of auxiliary storage slots on DASD are added to complement the new information on frames and page data set slots for SCM data blocks.

zEC12 – Flash Memory & Pagegable Large Pages...

⇨ New SCM statistics RMF Postprocessor Page Data Set Activity report

```
                                P A G E   D A T A   S E T   A C T I V I T Y

          z/OS V1R13            SYSTEM ID TRX2            DATE 03/10/2012            INTERVAL 15.00.012
                                                          TIME 13.00.00              CYCLE 1.000 SECONDS


NUMBER OF SAMPLES =     900                 PAGE DATA SET AND SCM USAGE
                                            --------------------------
PAGE                                                                  %    PAGE                  V
SPACE    VOLUME  DEV  DEVICE    SLOTS  ---- SLOTS USED ---  BAD      IN   TRANS NUMBER  PAGES  I
TYPE     SERIAL  NUM  TYPE      ALLOC    MIN    MAX   AVG   SLOTS    USE  TIME  IO REQ  XFER'D O  DATA SET NAME
PLPA     TRX2PP  D406 33903     71999  16851  16851  16851      0   0.00  0.000      0       0    PAGE.VTRX2PP.PLPA
COMMON   TRX2PP  D406 33903     35999     34     34     34      0   0.00  0.000      0       0    PAGE.VTRX2PP.COMMON
LOCAL    TRX2P1  D506 33903    593999      0      0      0      0   0.00  0.000      0       0 Y  PAGE.VTRX2P1.LOCAL1
SCM      N/A     N/A  N/A      131072  43151  43151  43151      0   0.00  0.000      0       0    N/A
```

System wide statistics of
4K SCM paging block usage and
SCM paging activity

- The Postprocessor Page Data Set Activity report displays SCM Paging Block information for page data set of type SCM.

zEC12 – Flash Memory & Pagegable Large Pages...

- Statistics for Fixed Large Memory Objects in RMF Postprocessor Virtual Storage Activity report

```
                              V I R T U A L   S T O R A G E   A C T I V I T Y
                                                                                PAGE    4
        z/OS V2R1                SYSTEM ID SYS3            DATE 11/25/2011          INTERVAL 14.59.996
                                 RPT VERSION V2R1 RMF      TIME 05.30.00            CYCLE 1.000 SECONDS

                                            PRIVATE AREA DETAIL

  JOB NAME -  JES2       MEMORY LIMIT - 20000M
  ...

  MEMORY ALLOCATION IN HIGH VIRTUAL MEMORY (ABOVE 2GB)

  BYTES                MIN              MAX              AVG    PEAK
    PRIVATE        1.823T 05.31.52 22.41T 05.50.36 11.51T  131.8T
    SHARED         485.1M 05.31.52  1822M 05.58.15 552.2M  1.333T
    COMMON         885.1M 05.31.52  1.22M 05.56.15   1.2T  1.927T

  MEMORY OBJECTS
    PRIVATE            80 05.31.52    160 05.35.52    110
    SHARED             20 05.31.52     70 05.35.52     23
    COMMON             30 05.31.52     60 05.35.52     44
    1 MB               20 05.31.52     70 05.35.52     25

  FRAMES (1 MB)
    FIXED               4 05.31.52     26 05.35.52     21
    PAGEABLE            4 05.31.52     26 05.35.52     21
```

Decomposition into
Fixed and Pageable Large Pages

- The Private Area Detail section of the RMF Postprocessor VSTOR report displays the MIN, MAX, and AVG values for the number of 1 MB frames that are used by FIXED and PAGEABLE memory objects that are owned by the reported address space.
- The PAGEABLE value also includes 1 MB frames that are used by DREF memory objects.

zEC12 – Flash Memory & Pagegable Large Pages...

New Large Page statistics in RMF Monitor III STORM report

System wide usage of Large Frame Area (Fixed Frames) and Pageable Large Frames

```
                    RMF V1R13  Storage Memory Objects

Samples: 60      System: TRX1  Date: 04/18/12  Time: 10.56.00

------------------------------- System Summary -------------------------------
 ---MemObj---    ---Frames---    -1MB MemObj-    --1MB Fixed--   -1MB Pageable-
 Shared     1    Shared   206    Total      4    Total     200   Initial    560
 Common    51    Common 24554    Common     3    Common     76   Dynamic      0
                 %Used    8.3                    %Used    52.8   %Used     10.5
------------------------------------------------------------------------------
           Service       ---- Memory Objects --- -1MB Frames- ----- Bytes -----
Jobname  C Class     ASID Total  Comm   Shr  1 MB Fixed Pgable Total  Comm   Shr

OMVS     S SYSTEM    0007    15     0     0     0     0     12 1351M     0     0
COMM641M B JESLOW    0035     3     1     0     1    30      0 32.0M 30.0M     0
STOR641M B JESLOW    0054     3     0     0     1    30      0 32.0M     0     0
```

Average number of 1 MB frames that are used by pageable and DREF memory objects owned by this address space

New and changed System Summary fields in Monitor III Storage Memory Objects (STORM) report:

Frames	%Used	Percentage of high virtual common storage used by the system.
1MB MemObj	Common	Average number of fixed memory objects that are allocated in high virtual common storage and can be backed in 1 MB frames. This value includes the memory objects that cannot be attributed to an address space.
1MB Fixed	Total	Total number of 1 MB frames that can be used by fixed memory objects. This value is equal to the size of the Large Frame Area in megabytes.
	Common	Average number of 1 MB high virtual common memory pages fixed in central storage. This value includes the pages that cannot be attributed to an address space or have not been freed during address space termination.
	%Used	Percentage of 1 MB frames that are used by fixed memory objects in the Large Frame Area regardless of whether the frames are actually used for 1 MB pages or used to satisfy 4K page requests on a constraint system.
1MB Pgable	Initial	Number of 1 MB frames that can be used by pageable and DREF memory objects. This value is calculated by the system at IPL time.
	Dynamic	Average number of 1 MB frames in the LFAREA that were used to satisfy 1 MB pageable page requests.
	%Used	Percentage of 1 MB frames that are used by pageable and DREF memory objects, regardless of whether the frames are actually used for 1 MB pages or used to satisfy 4K space requests on a constrained system.

New and changed Address Space related fields:

1MB Frames	Fixed	Average number of 1 MB frames in the Large Frame Area owned by this address space. Frames that are used to satisfy 4K space requests on a constrained system are not included.
	Pgable	Average number of 1 MB frames that are used by pageable and DREF memory objects owned by this address space. Pageable memory objects that have been fixed after allocation, are also included. Frames that are either used by common 1 MB pages or to satisfy 4K space requests on a constrained system are not included.

zEC12 – Coupling Facility Flash Memory

OA40515

- Coupling Facility Exploitation of SCM
  - Allows migration of MQSeries Shared Queue objects to flash memory when structure utilization exceeds threshold
  - MQ objects fetched back to real CF Storage when requested
  - Provides overflow capability for MQ Shared Queues to handle workload peaks
- New CF SCM statistics in
  - RMF Postprocessor Coupling Facility Activity (CF) report
  - RMF Monitor III Coupling Facility Overview (CFOVER) report
  - RMF Monitor III Coupling Facility Activity (CFACT) report

- Currently, the CF is a pure “real memory” system – all CF structures are allocated and backed entirely by real memory in the CF image. There is no storage hierarchy, hence no paging, no virtual storage, no disk I/O at all.
- With CFCC Level 19 the CF exploits internal flash memory for the purpose of migrating structure objects out to flash memory when the number of objects exceeds a calculated threshold and fetching them back into main CF storage when requested.
- Requires z/OS 2.1, or z/OS 1.13 with new function APAR OA40747
- Initially the CF Flash exploitation is targeted for MQ shared queues application structures.
- With RMF new function APAR OA40515, RMF provides the support for CFCC Storage Class Memory exploitation on the IBM zEnterprise EC12 (zEC12) server on z/OS 2.1 and z/OS 1.13.
- The support enhances RMF Postprocessor and Monitor III reports with various new CF SCM statistics.

zEC12 – Coupling Facility Flash Memory…

RMF Postprocessor Coupling Facility Activity Report:

- SCM statistics in new SCM STRUCTURE SUMMARY section

SCM and 'augmented' space capacity/usage in bytes

SCM I/O activity:
- Number of requests
- Number of of bytes transferred from/to SCM
- Average service time
- Standard deviation of service time

```
SCM STRUCTURE SUMMARY
-----------------------------------------------------------------------------------------------------------------------------------

                                 SCM SPACE  AUGMENTED  LST ENTRY LST ELEM  --- SCM READ ---- --- SCM WRITE ---  SCM AUX
     STRUCTURE                   MAX/       EST.MAX/   EST.MAX/  EST.MAX/  CNT/BYTE AVG ST/  CNT/BYTE AVG ST/  ENABLED
TYPE NAME                   ALG  %USED      %USED      CUR       CUR       X'FERRED STD_DEV  X'FERRED STD_DEV  CMD/%ALL

LIST THRLSTSCMKP1_2         KP1    16384M       312M     3932K    62915K     8332   1317.1     6830   1817.1       0
                                    33.2      10.3      1306K    20901K     8737M   154.3     7162M  1001.0     0.0
```

SCM statistics on CF structure level

SCM capacity/usage in number of list entries/elements

- The Postprocessor Coupling Facility Activity report is enhanced by a new 'SCM STRUCTURE SUMMARY' subsection as part of the 'USAGE SUMMARY' section.
- Per coupling facility, it lists all CF structures that have SCM assigned and shows structure related SCM maxima, percentage and usage counts:

ALG	Type of algorithm that is used by the coupling facility to control the movement of structure objects between coupling facility real storage and storage class memory:	
	KP1	KeyPriority1
	UNK	Unknown
SCM SPACE	MAX	Maximum amount of storage class memory that this structure can use (in bytes).
	%USED	Percentage of maximum amount of storage class memory that is in use by this structure.
AUGMENTED	Augmented space is additional CF real storage required to use SCM	
	EST.MAX	Estimated maximum amount of CF space that may be assigned as augmented space for this structure (in bytes).
	%USED	Percentage of maximum augmented space that is in use by this structure.
LST ENTRY	EST.MAX	Estimated maximum number of list entries that may reside in SCM for this structure.
	CUR	Number of existing structure list entries that reside in SCM.
LST ELEM	EST.MAX	Estimated maximum number of list elements that may reside in SCM for this structure.
	CUR	Number of existing structure list elements that reside in SCM.
SCM READ	CNT	The number of read operations against SCM that were either initiated - by a reference to list structure objects residing in SCM, or - as a prefetch operation in order to retrieve list structure objects in SCM that are expected to be referenced.
	BYTE X'FERRED	SCM read bytes transferred. This is the number of bytes transferred from SCM to CF.
	AVG ST	Average service time per read operation from SCM in microseconds.
	STD_DEV	Standard deviation of the service time for SCM read operations from SCM in microseconds.
SCM WRITE	CNT	The number of list write operations operations performed to SCM.
	BYTE X'FERRED	SCM write bytes transferred. This is the number of bytes transferred from CF to SCM.
	AVG ST	Average service time per write operation to SCM in microseconds.
	STD_DEV	Standard deviation of the service time for write operations to SCM in microseconds.
SCM AUX ENABLED	CMD	SCM auxiliary enabled command count. This is the number of commands that required the use of CF auxiliary frames.
	%ALL	Percentage of the SCM auxiliary enabled command count in relation to all requests for this structure.

zEC12 – Coupling Facility Flash Memory…

RMF Postprocessor Coupling Facility Activity Report:

- SCM statistics in STORAGE SUMMARY section

System wide SCM statistics

Augmented space:
- Total allocated in bytes
- % of CF storage

```
-------------------------------------------------------------------------------------------
 STORAGE SUMMARY
-------------------------------------------------------------------------------------------
                                          ALLOC      % OF CF      ------- DUMP SPACE -------
                                          SIZE       STORAGE     % IN USE   MAX % REQUESTED

TOTAL CF STORAGE USED BY STRUCTURES        3344M        3.3
TOTAL CF DUMP STORAGE                         6M        0.0          0.0                0.0
TOTAL CF AUGMENTED SPACE                     32M        0.0
TOTAL CF STORAGE AVAILABLE                98536M       96.7
                                          -------
TOTAL CF STORAGE SIZE                    101918M

                                          ALLOC      % ALLOCATED
                                          SIZE

TOTAL CONTROL STORAGE DEFINED            101918M        3.3
TOTAL DATA STORAGE DEFINED                    0K        0.0

                                       ASSIGNED      % IN USE     SUM MAX SCM

TOTAL CF STORAGE CLASS MEMORY            524288M        1.0           16384M
```

CF SCM space:
- Assigned capacity in bytes
- Capacity used
- Max possible SCM usage derived from current CFRM policy

- The 'STORAGE SUMMARY' section of the RMF Postprocessor Coupling Facility report is enhanced by SCM statistics about SCM space and augmented space usage.

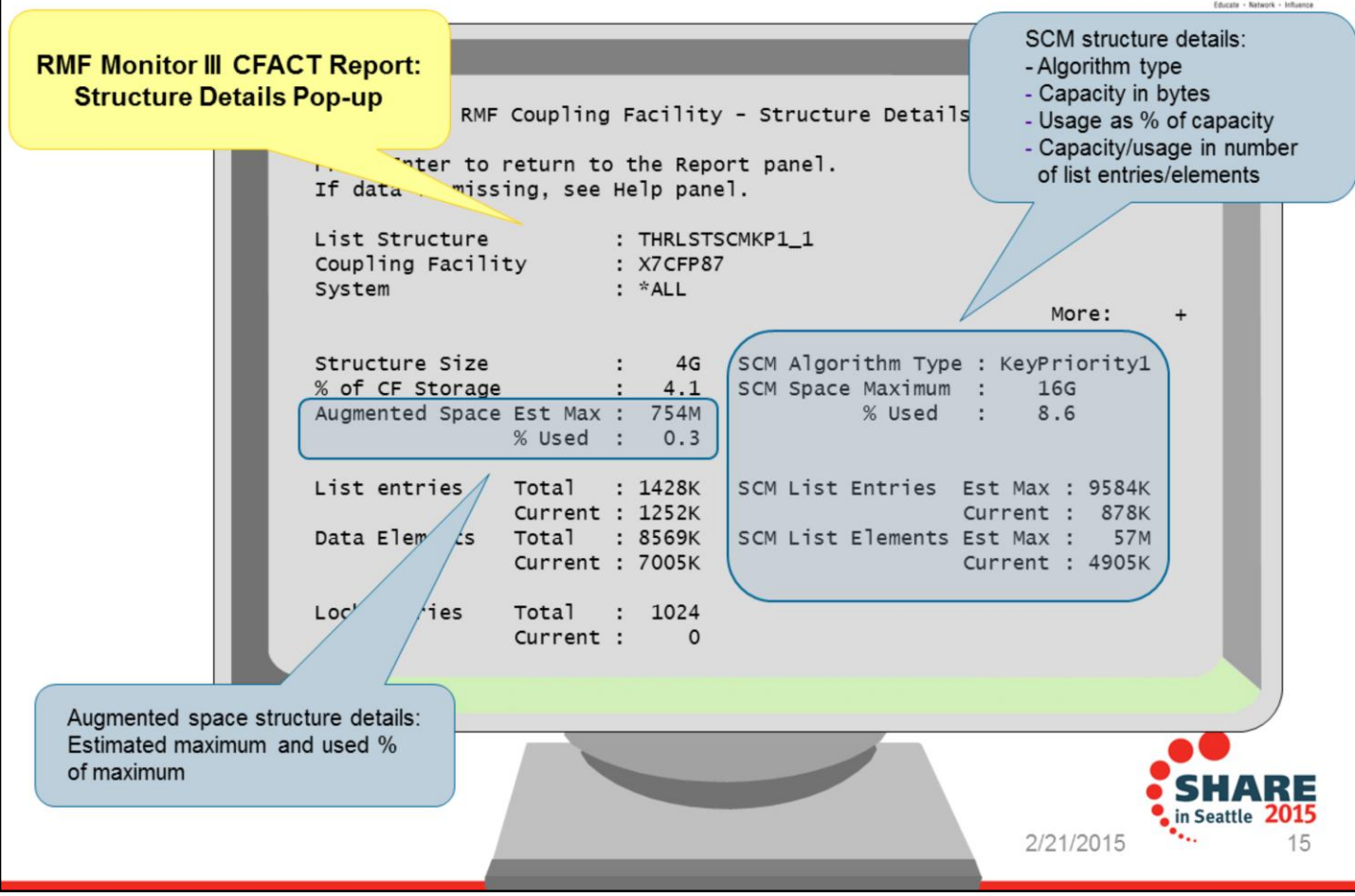
TOTAL CF AUGMENTED SPACE:

ALLOC SIZE	Total amount of CF storage used by all structures as augmented space (in bytes).
% OF CF STORAGE	Percentage of CF storage used by all structures as augmented space.

TOTAL CF STORAGE CLASS MEMORY:

ASSIGNED	Total CF SCM. This is the amount of SCM that may be concurrently used as structure extensions. Storage is assigned in increments of 4K bytes.
% IN USE	Percentage of SCM that is in use by all structures of the coupling facility.
SUM MAX SCM	Sum of the SCM maxima defined for all structures of the coupling facility.

zEC12 – Coupling Facility Flash Memory…



- With the RMF CF SCM support the Monitor III CFACT structure details pop-up window now provides the percentage of CF storage used by structure in new field:

% of CF Storage	The percentage of the total coupling facility storage allocated to this structure.

- For LIST structures that can use SCM storage extension the details pop-up window also shows SCM related statistics:

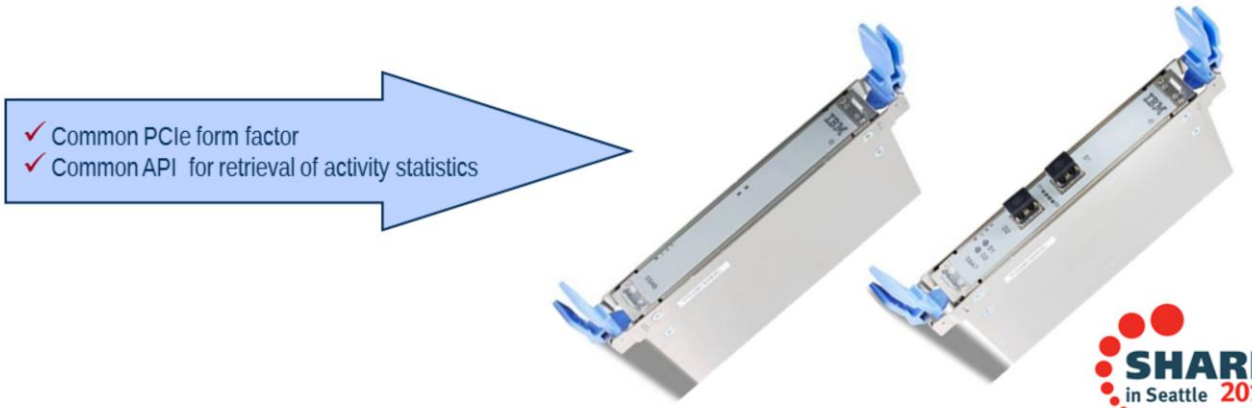
Augmented Space	Est Max	Estimated maximum amount of augmented space in bytes that can be assigned for this structure.
	% Used	Percentage of maximum augmented space that is in use by the structure.
SCM Space	Maximum	Maximum amount of SCM in bytes that this structure can use.
	% Used	Percentage of maximum SCM space that is in use by the structure.
SCM List Entries	Est Max	Estimated maximum number of list entries that can reside in SCM for the structure.
	Current	Number of existing structure list entries that reside in SCM.
SCM List Elements	Est Max	Estimated maximum number of list elements that can reside in SCM for the structure.
	Current	Number of existing structure list elements that reside in SCM.

zEC12 - Data Compression Express

- The z Enterprise Data Compression (zEDC) Express offering provides a low-cost data compression to z/OS system services and applications.
- Compression can be requested on
  - Dataset level via COMPACTION option in the SMS data class
  - System level via COMPRESS parameter in SYS1.PARMLIB(IGDSMSxx)
- Exploiters will see the following benefits
  - Reduced disk space
  - Increased performance for reading and writing compressed data
- The zEDC is implemented as a Peripheral Component Interconnect Express (PCIe) device that can be installed on zEC12 GA2 and zBC12.
- The compression function is provided via FPGA firmware. Other functions (also denoted as personality) may follow.
- You can install up to eight devices in a single machine where each device is sharable by up to 15 LPARs.

zEC12 - Shared Memory Communication

- Shared Memory Communication via Remote Direct Memory Access (SMC-RDMA or SMC-R) is a zEC12 feature that provides high performance CPC to CPC communication
- SMC-R actually offers the benefits of HiperSockets across processor boundaries. It takes advantage of high speed protocols and direct memory placement of data.
- SMC-R is totally transparent to applications.
- SMC-R is implemented as a Peripheral Component Interconnect Express (PCIe) device, also denoted RoCE adapter card (RDMA over Converged Ethernet)



zEC12 - Obtain PCIe Information by Command

- Use the console command D PCIE for general status information

PCIe Started Task

```
D PCIE
IQP022I 12.06.02 DISPLAY PCIE 970
PCIE     0013 ACTIVE
PFID  DEVICE TYPE NAME          STATUS  ASID  JOBNAME  PCHID VFN
0001  Hardware Accelerator      ALLC    0014  FPGHWAM  0380  0001
0020  10GbE RoCE                ALLC    00DE  VTAM     038C
0021  10GbE RoCE                CNFG                   03B0
0011  Hardware Accelerator      ALLC    0014  FPGHWAM  05C4  0001
```

PCIe Function ID

Owner Name

Physical Channel Identifier

Virtual Function Number

- Use the console command D PCIE,PFID=xxx for adapter details

```
D PCIE,PFID=001
IQP024I 12.10.53 DISPLAY PCIE 521
PCIE     0013 ACTIVE
PFID  DEVICE TYPE NAME          STATUS  ASID  JOBNAME  PCHID VFN
0001  Hardware Accelerator      ALLC    0014  FPGHWAM  0380  0001
 CLIENT ASIDS: NONE
 Application Description: zEDC Express
 Device State: Ready
 Adapter Info - Relid: 000000  Arch Level: 03
                Build Date: 06/28/2013  Build Count: 03
```

zEC12 - RMF Postprocessor PCIE Activity Report

- RMF Monitor III Data Gatherer collects PCIe performance statistics frequently and writes new SMF Record Type 74 Subtype 9
- The new RMF Postprocessor PCIE Activity Report provides detailed information about PCIE Express based functions.
  Currently supported functions are:
  - z Enterprise Data Compression (zEDC)
  - Shared Memory Communication via RDMA (SMC-R)
- New suboption PCIE | NOPCIE for Postprocessor REPORTS option
- Only available in XML format

RMF Postprocessor Duration Report [System SYSB] : PCIE Activity Report

Basic PCIe Metrics,e.g. PCI Load/Store and DMA Operations

RMF Version : z/OS V2R1 **SMF Data :** z/OS V2R1
Start : 11/14/2013-17.45.00 **End :** 11/14/2013-18.00.00 **Interval :** 000:15:00 hours

RoCE

General PCIE Activity

Function ID	Function PCHID	Function Name	Function Type	Function Status	Owner Job Name	Owner Address Space ID	Function Allocation Time	PCI Load Operations Rate	PCI Store Operations Rate	PCI Store Block Operations Rate	Refresh PCI Translations Operations Rate	DMA Address Space Count	DMA Read Data Transfer Rate	DMA Write Data Transfer Rate
0010	03BC	10GbE RoCE	15B31003	Allocated	VTAM	002A	900	0.921	9499	0	0.104	1	1993	21.8
0040	05EC	10GbE RoCE	15B31003	Allocated	VTAM	002A	900	0.923	9544	0	0.104	1	3434	12.6

- The PCIE performance data is collected by RMF Monitor III and stored in the new SMF 72 subtype 9.
- There is no RMF data gathering option. Data gathering is controlled by the SMF 72-9 setting in the active SMF parmlib member ERBSMFxx. Use TYPE/NOTYPE in the SMFPRMxx parmlib member to enable/disable the data gathering.
- Specify suboption PCIE in the RMF Postprocessor REPORTS control statement to create the new Postprocessor PCIE Activity Report.
- The PCIE Activity Report is only available in XML format.
- XML report can be created and displayed
  - via RMF Data Portal (SMF data from SMF buffer only)
  - via Spreadsheet Reporter (any SMF source)
- There is another option to display the SDELAY report in a web browser:
  - Download the XML output created by the RMF Postprocssor JOB to the toolkit directory provided with the RMF Postprocessor XML Toolkit
  - Open the report in the toolkit directory with your web browser.
- The PCIE Activity Report has up to three sections
  - General PCIE Activity section
    - Measurements for all PCIe functions
    - Data rates for PCI operations transferring data blocks from z/OS to the PCIe function
    - Direct memory access (DMA) read/write counters
  - Hardware Accelerator Activity section
    - Common accelerator request statisics: Queue Time, Execution Time and Request Rate
  - Hardware Accelerator Compression Activity
    - Measurements for zEDC
    - Compression related statistics:
      - Number of compressed bytes in and out
      - Number of decompressed bytes in an out
      - Buffer usage

zEC12 - RMF Postprocessor PCIE Activity Report

General PCIE Activity

Function ID	Function PCHID	Function Name	Function Type	Function Status	Owner Job Name	Owner Address Space ID	Function Allocation Time	PCI Load Operations Rate	PCI Store Operations Rate	PCI Store Block Operations Rate	Refresh PCI Translations Operations Rate	DMA Address Space Count	DMA Read Data Transfer Rate	DMA Write Data Transfer Rate
00A0	03C0	Hardware Accelerator	1014044B	Allocated	FPGHWAM	0012	14400	0	0.054	0	<0.01	1	21.8	21.2
00B0	0500	Hardware Accelerator	1014044B	Allocated	FPGHWAM	0012	14400	0	0.054	0	<0.01	1	1.19	1.19

zEDC

Common Request Statistics across all Personalities (Compression and future Personalities)

Hardware Accelerator Activity

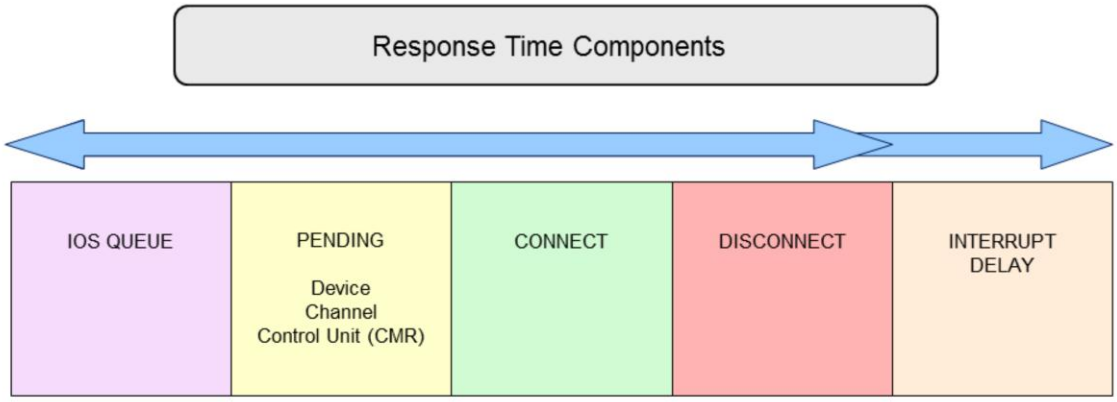
Function ID	Time Busy %	Request Execution Time	Std Dev for Request Execution Time	Request Queue Time	Std Dev for Request Queue Time	Request Size	Transfer Rate Total
00A0	0.002	27.7	4.77	462	70.0	69.6	0.056
00B0	0.002	28.6	5.07	487	39.8	69.6	0.059

Compression related Statistics

Hardware Accelerator Compression Activity

Function ID	Compression Request Rate	Compression Throughput	Compression Ratio	Decompression Request Rate	Decompression Throughput	Decompression Ratio	Buffer Pool Size	Buffer Pool Utilization
00A0	0.140	0.008	5.03	0.665	0.008	0.198	15360	0
00B0	0.139	0.008	5.03	0.713	0.008	0.198	15360	0

zEC12 – I/O Interrupt Delay Time



- Interrupt Delay time measures the time between when the I/O completes and z/OS issues the TSCH instruction to retrieve the results.
- How long does it take for z/OS to see and process the interrupt after I/O completes ?

- The response time that is reported for a device is the sum of IOS QUEUE time, PENDING time, CONNECT time and DISCONNECT time.
- However, there is another response time component that was not measured and reported but impacts the overall I/O response time as seen by the application:
  - The I/O Interrupt Delay time
- Interrupt Delay time is the time between when the I/O completes and z/OS issues the TSCH instruction to retrieve the results. In other words:
  How long does it take for z/OS to see and process the interrupt after I/O completes ?

zEC12 – I/O Interrupt Delay Time

- I/O interrupt delay time (aka I/O elongation) occurs when an I/O is delayed due to
  - a non-dispatched LPAR
  - the lack of an interrupt enabled CP
- I/O interrupt delay time is NOT included in the I/O response time
- I/O interrupt delay time could NOT be measured in the past
- High ratios of interrupts handled via TPI (Test Pending Interrupt) are an indicator for significant I/O elongation

CPU Activity

CPU : 2817 Model : 729 H/W Model : M32 Sequence Code : 00000000000E3206 HiperDispatch : NO
CPC Capacity : 2780 Change Reason : NONE

CPU Number	CPU Type	Time% Online	Time% LPAR Busy	Time% MVS Busy	Time% Parked	LOG PROC Share%	HiperDispatch Priority	I/O Interrupts Rate	I/O Interrupts% via TPI
0	CP	100.00	1.60	1.57	------	31.9		1.87	0.18
1	CP	100.00	1.14	1.12	------	31.9		2.69	0.33
2	CP	100.00	0.85	0.83	------	31.9		2.72	0.29
TOTAL/AVERAGE	CP		1.20	1.18		95.7		7.27	0.28

- There are two possible reasons for an I/O interrupt delay:
  - The LPAR is not dispatched
  - There are not enough CPs that are enabled to handle interrupts
- I/O interrupt delay time is NOT included in the total I/O response time as reported by RMF
- In the past I/O interrupt delay time could not even been measured.
- When data transfer is complete on an I/O device, the channel subsystem will attempt to find a CP (idle or busy) that is enabled for interrupts
- After the CP completes processing for an I/O interrupt, it issues a Test Pending Interrupt (TPI) instruction to determine whether there are any further I/O interrupts pending. If yes, the processor proceeds to service that interrupt
- Hence, an interrupt handled via TPI is a queued interrupt, that could not be handled instantly by an alternate I/O enabled processor
- Not more than 20-30% should be handled via TPI, otherwise it is time to examine the I/O interrupt delay time

zEC12 – I/O Interrupt Delay Time

- New field AVG INT DLY in RMF Postprocessor Device Activity report.
- Caution: Interrupt Delay Time is not included in AVG RESP Time

OA39993

```
          D I R E C T   A C C E S S   D E V I C E   A C T I V I T Y

       z/OS V1R13            SYSTEM ID TRX2          DATE 11/23/2011        INTERVAL 14.59.998
                             RPT VERSION V1R13 RMF   TIME 12.30.00          CYCLE 1.000 SECONDS

TOTAL SAMPLES =    900   IODF = 01   CR-DATE: 11/15/2011   CR-TIME: 07.33.54     ACT: POR
                                                  DEVICE   AVG  AVG   AVG  AVG  AVG   AVG  AVG  AVG
STORAGE  DEV  DEVICE   NUMBER  VOLUME PAV  LCU  ACTIVITY RESP IOSQ  CMR  DB   INT   PEND DISC CONN
 GROUP   NUM  TYPE     OF CYL  SERIAL           RATE     TIME TIME  DLY  DLY  DLY   TIME TIME TIME
XTEST    2208 33903      3339  TRXSX9   1  0032   0.001  .384 .000  .128 .000 .123  .256 .000 .128
XTEST    2209 33903      3339  TRXSXA   1  0032   0.001  .256 .000  .000 .000 .135  .256 .000 .000
         220A 33909     10017  TRXT01   1  0032   0.000  .000 .000  .000 .000 .000  .000 .000 .000
         220B 33909     10017  TRXT02   1  0032   0.000  .000 .000  .000 .000 .000  .000 .000 .000
```


- V1.13 RMF is extended to report the interrupt delay time measurement for devices attached to the IBM zEnterprise EC12 server.
- Support available with new function APAR OA39993
- New Interrupt Delay Time reported in RMF Postprocessor Device Activity report:

- AVG INT DLY: The average interrupt delay time in units of milliseconds encountered for I/O requests to this device.
  For each I/O request, the time is measured from when the I/O operation is complete to when the operating system begins to process the status.

zEC12 – Enhanced CF Link Reporting

- In a System z environment different types of coupling links can be used to connect a Coupling Facility (CF) to the operating system.
- Each coupling link type has effect on link performance, response times and coupling overheads.
- For configurations covering large distances, the time spent on the link can be the largest part of the response time.

OA37826

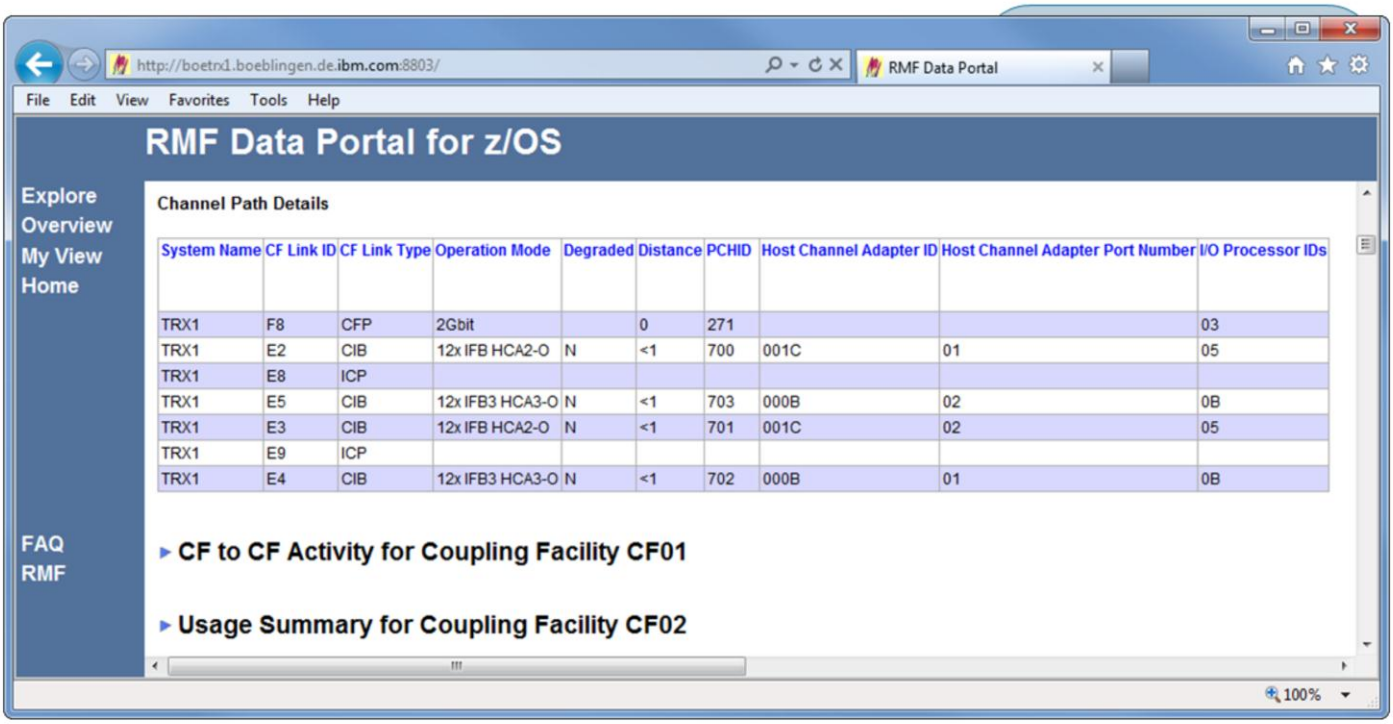
Type	Description	CHPID Type	Link Rate	Distance
ISC-3	InterSystem Channel-3	CFP (CF Peer)	2 Gbps	10 km unrepeated (6.2 miles) 100 km repeated
PSIFB	12x IB-DDR Parallel Sysplex InfiniBand 12x IB-SDR Parallel Sysplex InfiniBand	CIB (Coupling over Infiniband)	6 GBps 3 GBps (When connected to a System z9 EC or System z9 BC)	150 meters (492 feet)
PSIFB-LR	1x IB-SDR Parallel Sysplex InfiniBand – Long Reach Double data rate (1 x IB-DDR) is supported if connected to a System z qualified DWDM supporting DDR.	CIB (Coupling over Infiniband)	2.5 Gbps 5.0 Gbps	10 km unrepeated (6.2.miles) 100 km repeated
IC	Internal Coupling Channel	ICP (Internal CF Peer)	Internal speeds	n/a

- With system z, different types of coupling links can be used to connect a Coupling Facility (CF) to an operating system. Each coupling link type has effect on link performance, response times and coupling overheads.
- For configurations covering large distances, the time spent on the link can be the largest part of the response time. The table gives an overview of the current types of coupling links supported by system z.
- With new function APAR OA37826 the Postprocessor Coupling Facility Activity report and Monitor III CFSYS report are enhanced to display additional channel path information for channel paths of type CIB or CFP.
- The RMF support is available for z/OS V1.12 and V1.13 RMF

zEC12 – Enhanced CF Link Reporting

RMF Postprocessor Coupling Facility Activity Report:

⇨ New CHANNEL PATH DETAILS in Subchannel Activity Section



System Name	CF Link ID	CF Link Type	Operation Mode	Degraded	Distance	PCHID	Host Channel Adapter ID	Host Channel Adapter Port Number	I/O Processor IDs
TRX1	F8	CFP	2Gbit		0	271			03
TRX1	E2	CIB	12x IFB HCA2-O	N	<1	700	001C	01	05
TRX1	E8	ICP							
TRX1	E5	CIB	12x IFB3 HCA3-O	N	<1	703	000B	02	0B
TRX1	E3	CIB	12x IFB HCA2-O	N	<1	701	001C	02	05
TRX1	E9	ICP							
TRX1	E4	CIB	12x IFB3 HCA3-O	N	<1	702	000B	01	0B

- New function APAR OA37826 adds a new CHANNEL PATH DETAILS section to the Subchannel Activity and CF To CF Activity sections of the RMF Postprocessor Coupling Facility Activity report.
- The enhanced subchannel reporting function is provided by the zEnterprise EC12 (zEC12) processor for CF link types CIB or CFP and require the following prerequisites:
  - To report or display the additional channel path information for directly-connected CFs, z/OS must be running on a zEC12 processor.
  - To report or display the additional channel path information for remotely-connected CFs, the directly-connected CF must be running on a zEC12 processor (CFLEVEL 18 or higher).
- For a full exploitation of this new support in a sysplex environment it is necessary to install this APAR on all systems of the sysplex.

OPERATION MODE: Channel path operation mode. It describes the data rate, bandwidth, protocol and adapter type of the channel path.

Data rates for CFP channels paths:
- 1GBIT - 1.0625 Gbit/sec
- 2GBIT - 2.125 Gbit/sec

Bandwidth for CIB channels paths:
- 1x - single bandwidth
- 12x - twelve-fold bandwidth

Protocol for CIB channels paths:
- IFB or IFB3

Adapter types for CIB channels paths:
- HCA-O, HCA-O LR, HCA3-O or HCA3-O LR

Unknown operation mode:
- UNKNOWN

zEC12 – Enhanced CF Link Reporting…

RMF Postprocessor Coupling Facility Activity Report:

- New CHANNEL PATH DETAILS in CF to CF Activity Section

```
COUPLING FACILITY NAME = X5CFP87
-------------------------------------------------------------------------------------------------
                                                CF TO CF ACTIVITY
-------------------------------------------------------------------------------------------------
          # REQ                          ----------- REQUESTS ----------   ----------
 PEER     TOTAL       -- CF LINKS --                 #   -SERVICE TIME(MIC)          #
 CF       AVG/SEC     TYPE       USE                 REQ     AVG    STD_DEV          REQ

 X5CFH89  243089      CFP          2       SYNC  243089    18.6       4.9    SYNC     0     0.0                 0.0       0.0
          270.1

 X5CFR89  34647K      CIB          6       SYNC   34647K   12.4       4.0    SYNC    83                 8.4       9.0       0.0
          38497
-------------------------------------------------------------------------------------------------------------------------------
                                                CHANNEL PATH DETAILS
-------------------------------------------------------------------------------------------------------------------------------
 PEER CF   ID  TYPE   OPERATION MODE      DEGRADED   DISTANCE

 X5CFH89   02  CFP    2GBIT                                 0
           03  CFP    2GBIT                                 0
 X5CFR89   D0  CIB    12X IFB  HCA3-O        N             <1
           D1  CIB    12X IFB  HCA3-O        N             <1
           D2  CIB    12X IFB  HCA3-O        N             <1
           D3  CIB    12X IFB  HCA3-O        N             <1
           E0  CIB    12X IFB3 HCA3-O        Y             <1
           E1  CIB    12X IFB3 HCA3-O        Y             <1
```

Channel path details for CF to CF links of type CIB or CFP:

- Operation mode
- Degraded status
- Link distance

DEGRADED: Character **Y** in this column indicates that the channel path is operating at reduced capacity (degraded) or not operating at all.

DISTANCE: Estimated distance in kilometer. The value is calculated as follows:

Average round-trip path time in microseconds

---
10 microseconds / kilometer

The value '<1' means that the estimated distance is below one kilometer.
A value of zero means that the time was not measured.

PCHID: Physical channel ID.

HCA ID: The hexadecimal host channel adapter ID.

HCA PORT: The hexadecimal host channel adapter port number.

IOP IDS: The hexadecimal identifiers of I/O processors (System Assist Processors) to which the channel path is accessible.

zEC12 – Enhanced CF Link Reporting…

RMF Monitor III CFSYS Report:
Channel Paths Details Pop-up

```
              RMF Coupling Facility - Subchannels and Paths

Press Enter to return to the Report panel.

Details for System      : TRX1
Coupling Facility       : CF01

Subchannels  Generated : 149
             In Use    :  49
             Max       :  49

Channel Path Details:
                                                  ---HCA---
ID Type  Operation Mode      Deg Distance PCHID   ID Port --IOP IDs--
                                                           More:     +
E2 CIB   12x IFB  HCA2-O      N        <1  0700 001C   01 05
E3 CIB   12x IFB  HCA2-O      N        <1  0701 001C   02 05
E4 CIB   12x IFB3 HCA3-O      N        <1  0702 000B   01 0B
E5 CIB   12x IFB3 HCA3-O      N        <1  0703 000B   02 0B
E8 ICP
E9 ICP
F8 CFP   2Gbit                          0  0271           03
```

Scrollable List
of all available Channel Paths
to Coupling Facility

- The 'Subchannels and Paths' pop-up panel of the RMF Monitor III Coupling Facility Systems Report (CFSYS) report is modified to show a scrollable list of all channel paths that are available to the reported coupling facility.

zEC12 - Support of Crypto Express4 Card

OA37016

```
                                C R Y P T O   H A R D W A R E   A C T I V I T Y

            z/OS V1R13                SYSTEM ID SYSF            DATE 11/29/2011
                                      RPT VERSION V1R13 RMF     TIME 16.00.00
-------- CRYPTOGRAPHIC COPROCESSOR --------
           -------- TOTAL --------  KEY-GEN
TYPE   ID   RATE  EXEC TIME  UTIL%    RATE
CEX2C   0   0.00    0.000     0.0     0.00
        1   2.16    295.9    63.9     2.14
CEX3C   2   0.00    0.000     0.0     0.00
CEX4C   4   2.15    227.8    48.9     2.15

-------- CRYPTOGRAPHIC ACCELERATOR ----------------------------------------------------------------------------------------
           -------- TOTAL -------          -- ME-FORMAT RSA OPERATIONS --  -- CRT-FORMAT RSA OPERATIONS --
TYPE   ID   RATE  EXEC TIME UTIL%    KEY     RATE     EXEC TIME    UTIL%     RATE     EXEC TIME    UTIL%
CEX2A   3  766.9    0.434    33.3   1024    362.4      0.521       18.9     369.5      0.183        6.8
                                    2048     0.00      0.000        0.0     34.99      2.175        7.6
CEX4A   5  998.9    0.365    36.5   1024    246.4      0.534       13.2     554.3      0.205       11.3
                                    2048     0.00      0.000        0.0     83.16      0.689        5.7
                                    4096     0.00      0.000        0.0     115.1      0.547        6.3

-------- ICSF SERVICES ----------------------------------------------------------------------------------------------------
         ---- ENCRYPTION -----    ---- DECRYPTION -----    ----- MAC ------    -------- HASH ----------    ------ PIN -------
          SDES    TDES     AES    SDES    TDES     AES    GENERATE  VERIFY    SHA-256  SHA-512    TRANSLATE   VERIFY
 RATE     0.00    0.00    0.00    0.00    0.00    0.00       0.00    0.00      0.00     0.00      0.00        0.00      0.00
 SIZE     0.00    0.00    0.00    0.00    0.00    0.00       0.00    0.00      0.00     0.00
```

All measurements available for Crypto Express4 Card

- RMF new function APAR OA37016 introduces the RMF support of the Crypto Express4S feature on the zEC12
- In detail, RMF collects and reports performance measurements for operations executed on Crypto Express4S co-processors (CEX4C) and accelerators (CEX4A).
- CEX4C measurements are stored in the Cryptographic Coprocessor data section and CEX4A measurements are stored in the Cryptographic Accelerator data section of SMF record 70 subtype 2.
- The Postprocessor Crypto Activity report provides CEX4C measurements in the CRYPTOGRAPHIC COPROCESSOR section and CEX4A measurements in the CRYPTOGRAPHIC ACCELERATOR section.
- The RMF support is available for z/OS V1.10 and above

RMF Enhancements at a Glance

- IBM zEnterprise EC12 Support
  - Statistics for Flash Memory and Pageable Large Pages
  - Support of Coupling Facility Flash Memory
  - z Enterprise Data Compression Express Reporting (zEDC)
  - Shared Memory Communication Reporting (SMC-R)
  - Support of I/O Interrupt Delay Time Facility
  - Extended Infiniband Link Reporting
  - Support of Crypto Express4 Card

- IBM z13 Support
  - RMF support for SMT, Crypto Express5S …
  - PCIe reporting enhancements

- Exploitation of System z Integrated Information Processors
  - zIIP Usage Option for Monitor III Gatherer

- Postprocessor XML formatted Reports
  - Transition to XML Format almost complete
  - Advanced Sorting and Filtering Capabilities

- RMF XP Enhancements
  - SMF Type 104 Recording for all Platforms

- z/OSMF Resource Monitoring
  - Context sensitive Application Linkage to WLM
  - Historical Reporting & Spreadsheet Export

RMF z13 Support - Overview

- RMF is enhanced to monitor performance of logical zIIP cores and their threads configured in a Simultaneous Multithreading (SMT) environment on z13. **OA44101**
- RMF provides new measurements for PCIe-attached RoCE and zEDC devices on z13: **OA44524**
  - Supports new z13 10GbE shared RoCE (SR-IOV) express feature
- RMF support for Crypto Express5S (CEX5) card and new ICSF service measurements: **OA43493**
  - RSA Digital Signature Generate and Verify callable services
  - ECC Digital Signature Generate and Verify callable services
  - AES MAC Generate and Verify callable services
  - FPE Encipher, Decipher and Translate callable services
- RMF support for LPARs with up to 4 TB real storage. **OA44503**
- RMF support for z13 IBM Integrated Coupling Adapter (ICA SR) that provides PCIe based short-distance coupling links of type CS5 **OA44502**

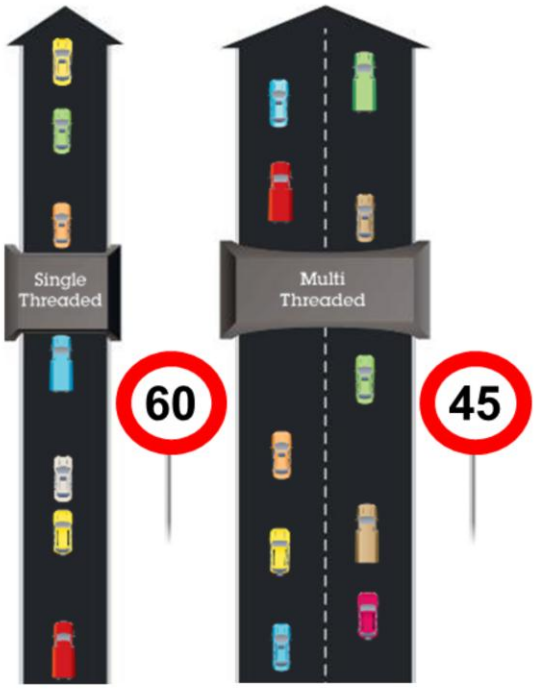
- RMF z13 Toleration support for z/OS 1.10 and 1.11 **OA45890**
- RMF z13 Toleration support for z/OS 1.12 and 1.13 **OA44833**

- With various new function APARs, RMF exploits the new functionality of the IBM z13:
  - OA44101: RMF support for the Simultaneous Multithreading (SMT) environment on z13 PTF available for z/OS 2.1
  - OA44524: RMF PCIE enhancements for RoCE and zEDC devices on z/OS 2.1.
  - OA43493: RMF support for the Crypto Express5S (CEX5) card and new ICSF service measurements. The support is available for z/OS 1.13 and z/OS 2.1.
  - OA44503: RMF support for z/OS 2.1 LPARs on z13 with up to 4TB real storage.
  - OA44502: RMF support for z13 IBM Integrated Coupling Adapter (ICA SR) that provides PCIe based short-distance coupling links of type CS5. PTFs available for z/OS 1.13 and z/OS 2.1.
- RMF tolereation support for IBM z13:
  - OA45890: z13 toleration for z/OS 1.10 and z/OS 1.11
  - OA44833: z13 toleration for z/OS 1.12 and z/OS 1.13

.

z13 - Simultaneous Multithreading (SMT)

- "Simultaneous multithreading (SMT) permits multiple independent threads of execution to better utilize the resources provided by modern processor architectures."*
- With z13, SMT allows up to two instructions per core to run simultaneously to get better overall throughput
- SMT is designed to make better use of processors
- On z/OS, SMT is available for zIIP processing:
  - ⇨ Two concurrent threads are available per core and can be turned on or off
  - ⇨ Capacity (throughput) usually increases
  - ⇨ Performance may in some cases be superior using single threading



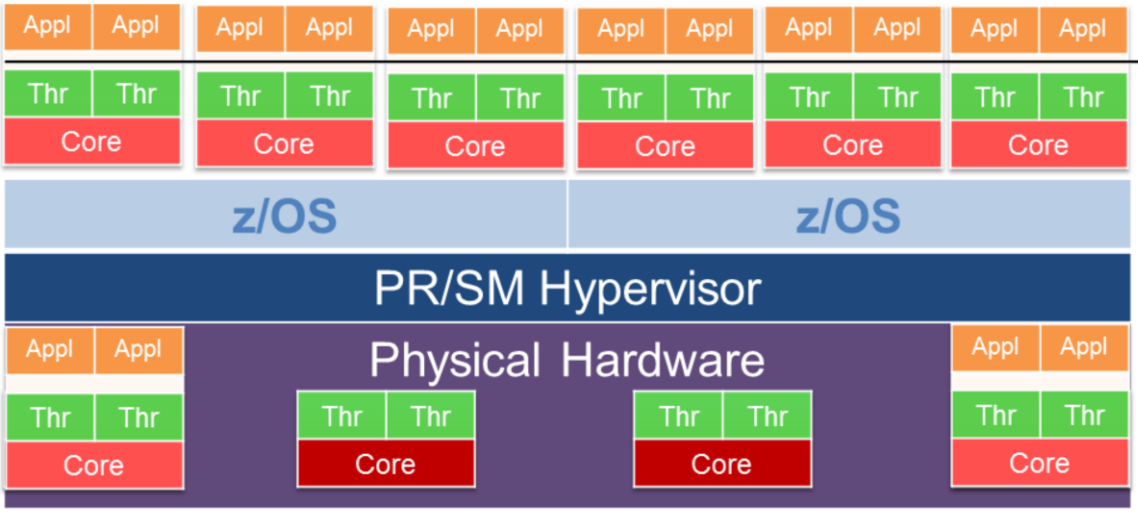
Two lanes process more traffic overall

Note: Speed limit signs for illustration only

*Wikipedia®

- Simultaneous multithreading (SMT) allows two active instruction streams (threads) per core, each dynamically sharing the core's execution resources. SMT will be available in IBM z13 for workloads running on the Integrated Facility for Linux (IFL) and the IBM z Integrated Information Processor (zIIP).
- SMT utilizes the core resources more efficiently: When a thread running on a core encounters a cache miss and can no longer make progress, the core switches to run a different thread that is ready to execute.
- Each thread runs slower than a non-SMT core, but the combined 'threads' throughput is higher. The overall throughput benefit depends on the workload.

z13 - SMT Exploitation



- SMT Aware OS informs PR/SM that it intends to exploit SMT
  - PR/SM can dispatch any OS core to any physical core
  - OS controls the whole core – must follow rules
    - Maximize core throughput (Drive cores with high Thread Density [2])
    - Maximize core availability (Meet workload goals using fewest cores)
- SMT is transparent to applications
- LOADxx and IEAOPTxx parmlib options to enable SMT on z/OS:
  - LOADxx: PROCVIEW **CORE|CPU**
  - IEAOPTxx: MT_ZIIP_MODE={1 | 2}

- The use of SMT mode can be enabled on an LPAR by LPAR basis via operating system (OS) parameters
- Once the OS switches to SMT mode, the only way back to single thread (ST) mode is via a disruptive action (re-activate the partition or re-IPL it).
- With the SMT enabled mode it is possible to dynamically switch between MT-1 (multi thread) and MT-2 mode for the processor types that support MT-2
- z/OS introduces new options for the LOADxx and IEAOPTxx parmlib members that are used to enable/disable SMT support and specify the MT mode of a processor class:
  - LOADxx parmlib option PROCVIEW CORE|CPU enables/disables SMT for the life of the IPL
    - PROCVIEW CORE on z13 enables SMT support
    - IPL required to switch between PROCVIEW CPU and CORE
  - New IEAOPTxx parameter to control the MT mode for zIIP processors
    - MT_ZIIP_MODE=1 specifies MT-1 mode for zIIPs (one active thread per online zIIP core)
    - MT_ZIIP_MODE=2 specifies MT-2 mode for zIIPs (two active threads per online zIIP core)
    - When PROCVIEW CPU is specified the processor class MT mode is always 1
    - SET OPT=xx operator command allows to switch dynamically between MT-1 and MT-2 mode
    - MT-2 mode requires HiperDispatch to be in effect
  - z/OS SMT Terminology:
- z/OS logical processor (CPU) ➔ Thread
  - A thread implements (most of) the System z processor architecture
  - z/OS dispatches work units on threads
  - In MT mode two threads are mapped to a logical core
- Processor core ➔ Core
- PR/SM dispatches logical core on a physical core
  - Thread density 1 (TD1) when only a single thread runs on a core
  - Thread density 2 (TD2) when both threads run on a core

z13 - z/OS SMT Metrics

- Capacity Factor (CF)
  - How much work core actually completes for a given workload mix at current utilization - relative to single thread
  - MT-1 Capacity Factor is 1.0 (100%)
  - MT-2 Capacity Factor is workload dependent
- Maximum Capacity Factor (mCF)
  - How much work a core can complete for a given workload mix at most
- Core Busy Time
  - Time any thread on the core is executing instructions when core is dispatched to physical core
- Average Thread Density
  - Average number of executing threads during **Core Busy Time** (Range: 1.0 - 2.0)
- Productivity
  - Core Busy Time Utilization (percentage of used capacity) for a given workload mix
  - Productivity represents capacity in use (**CF**) relative to capacity total (**mCF**) during **Core Busy Time**.
- Core Utilization
  - Capacity in use relative to capacity total over some time interval
  - Calculated as **Core Busy Time** x **Productivity**

Actual MT-2 Efficiency

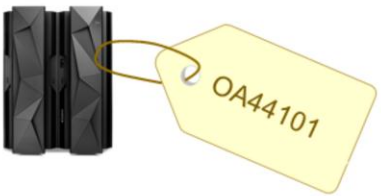
Estimated max MT-2 Efficiency

% Used MT-2 Core Capacity during Core Busy Time

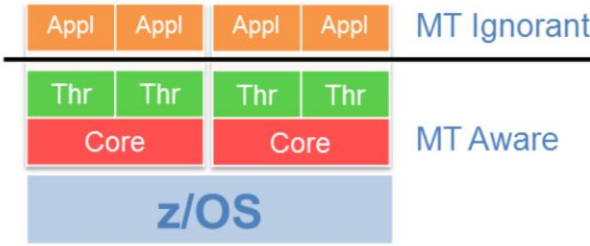
% Used MT-2 Core Capacity during Measurement Interval

- z/OS SMT introduces several new metrics to describe how efficiently the core resources could be utilized and how efficiently they are actually utilized.

z13 – RMF and SMT



- RMF enhanced with new metrics to monitor MT-2 efficiency and core utilization
- Reinterpret the meaning of exiting RMF metrics:
  - CPU metrics on core granularity (e.g. APPL%/EAPPL%)
  - CPU metrics on thread granularity (e.g. MVS BUSY%)
  - SMT updates in RMF Documentation
    - Enhanced metrics descriptions
    - General terminology:
      - „Processor“ ➔ logical Core
      - „logical Processor“ ➔ Thread
  - MT-1 Equivalent Time
    - z/OS CPU time consumed by work units (TCBs, SRBs) provided in terms of MT-1 equivalent time
    - Time it would have taken to run same work in MT-1 mode
    - Reflected in all RMF metrics reporting CPU consumption of workloads as CPU times or service units



- The RMF support for SMT provides new SMT related metrics to allow capacity planning and performance analysis in SMT environments.
- RMF supports SMT environments by extending the
  - Postprocessor CPU activity report
  - Monitor III CPC capacity report
  - Overview Conditions based on SMF 70.1
- RMF new function APAR OA44101 provides the SMT support for z/OS 2.1.
- The architecture introduced with SMT requires a reinterpretion of existing RMF metrics:
  - CPU metric data can now be on core or thread level granularity
  - z/OS charges CPU time consumed by work units (TCBs, SRBs) in terms of MT-1 equivalent time. The MT-1 equivalent time is the time it would have taken to run same work in MT-1 mode. All RMF metrics reporting CPU consumption of workloads as CPU time or service units reflect MT-1 equivalent time.

z13 – SMT: Postprocessor CPU Activity Report

- PP CPU activity report displayed in “old” format when SMT is inactive
- PP CPU activity report provides new metrics when SMT is active
  - MT Productivity and Utilization of each logical core
  - Multi-Threading Analysis section displays MT Mode, MT Capacity Factors and average Thread Density
- One data line in PP CPU activity report represents one thread (CPU)
  - CPU NUM designates the logical core
- Some metrics like TIME % ONLINE and LPAR BUSY provided at core granularity only

```
                                        C P U   A C T I V I T Y
z/OS V2R1                 SYSTEM ID CB8B              DATE 02/02/2015            INTERVAL 15.00.004
                          RPT VERSION V2R1 RMF        TIME 11.00.00              CYCLE 1.000 SECONDS
---CPU---     --------------- TIME % ------------    --- MT % ----  LOG PROC      --I/O INTERRUPTS-
NUM  TYPE     ONLINE    LPAR BUSY  MVS BUSY PARKED   PROD    UTIL   SHARE %       RATE      % VIA TPI
 0    CP      100.00    68.07      67.94      0.00   100.00  68.07  100.0  HIGH   370.1     13.90
 1    CP      100.00    46.78      46.78      0.00   100.00  46.78   52.9  MED     5.29     16.93
...
TOTAL/AVERAGE            8.66      54.17             100.00   8.66  152.9         375.3     13.95
 A    IIP     100.00    48.15      41.70      0.00    85.84  41.33  100.0  HIGH
                                   35.66      0.00
 B    IIP     100.00    38.50      32.81      0.00    85.94  33.09  100.0  HIGH
                                   26.47      0.00
...
TOTAL/AVERAGE           29.48      23.23              86.47  25.39  386.7
------------ MULTI-THREADING ANALYSIS ---------------
 CPU TYPE     MODE      MAX CF            CF       AVG TD
      CP        1       1.000          1.000       1.000
      IIP       2       1.485          1.279       1.576
```

MT-2 core capacity used:
MT Core Productivity
x TIME % LPAR BUSY

Productivity of logical core while dispatched to physical core

- The CPU Activity section reports on logical core and logical processor activity. For each processor, the report provides a set of calculations that are provided at a particular granularity that depends on whether multithreading is disabled (LOADxx PROCVIEW CPU parameter is in effect) or enabled (LOADxx PROCVIEW CORE parameter is in effect).
- If multithreading is disabled for a processor type, all calculations are at logical processor granularity.
- If multithreading is enabled for a processor type, some calculations are provided at logical core granularity and some are provided at logical processor (thread) granularity. The CPU Activity section displays exactly one report line per thread showing all calculations at logical processor granularity. Those calculations that are provided at core granularity are only shown in the same report line that shows the core id in the CPU NUM field and which is representing the first thread of a core.
- The following calculations are on a per logical processor basis when multithreading is disabled and on a per logical core basis when multithreading is enabled
  - Percentage of the interval time the processor was online
  - LPAR view of the processor utilization (LPAR Busy time percentage)
  - Percentage of a physical processor the logical processor is entitled to use
  - Multithreading core productivity (only reported when multithreading is enabled)
  - Multithreading core utilization (only reported when multithreading is enabled)
- The following calculations are on a per logical processor basis regardless whether multithreading is enabled or disabled:
  - MVS view of the processor utilization (MVS Busy time percentage)
  - Percentage of the online time the processor was parked (in HiperDispatch mode only)
  - I/O interrupts rate (general purpose processors only)
  - Percentage of I/O interrupts handled by the I/O supervisor without re-enabling (general purpose processors only)

z13 – SMT: Monitor III CPC Report

MT Mode and Productivity for zIIP processors

```
                          RMF V2R1   CPC Capacity                    Line 1 of 50

Samples: 60       System: CB88  Date: 02/02/15  Time: 11.00.00  Range: 60    Sec

Partition:   CB88        2964 Model 731
CPC Capacity:     3935   Weight % of Max: 50.1   4h Avg:  138   Group:   N/A
Image Capacity:   1777   WLM Capping %:    0.0   4h Max:  177   Limit:   N/A
MT Mode IIP:         2   Prod % IIP:      80.9
```

```
Partition  --- MSU ---  Cap  Proc    Logical Util %   - Physical Util % -
             Def   Act  Def   Num    Effect   Total   LPAR  Effect  Total

*CP                          390                       0.8    43.7   44.5
CB8B           0   192  NO  10.0      15.0    15.1    0.0     4.8    4.9
CB8D           0   134  NO  15.0       7.0     7.0    0.0     3.4    3.4
CB8E           0   330  NO  14.0      18.4    18.6    0.1     8.3    8.4
CB88           0   182  NO  14.0      10.2    10.3    0.0     4.6    4.6
C05            0   140  NO  14.0       7.9     7.9    0.0     3.5    3.6
C06            0   150  NO  14.0       8.4     8.4    0.0     3.8    3.8
LP1            0   507  NO   4.0       100     100    0.0    12.9   12.9
```

SMT mode enabled:
Processor data at logical core granularity

SMT mode disabled:
Processor data at logical processor granularity

- RMF Monitor III CPC report displays performance data for all partitions belonging to the CPC
- If multithreading is enabled the processor data is reported at logical core granularity, otherwise processor data is reported at logical processor granularity
- The report header is enhanced with the information about MT Mode and Productivity for the zIIP processors.
- Additional SMT metrics are available as hidden report header fields:
  - Multi-Threading Maximum Capacity Factor for IIP
  - Multi-Threading Capacity Factor for IIP
  - Average Thread Density for IIP
- These hidden report header fields can be displayed, if the CPC report is invoked in the RMF Data Portal for z/OS web browser frontend.

z13 – SMT: Postprocessor Workload Activity Report

```
                                W O R K L O A D   A C T I V I T Y

z/OS V2R1       SYSPLEX UTCPLXCB          DATE 02/02/2015     INTERVAL 15.00.004          MODE = GOAL
                RPT VERSION V2R1 RMF      TIME 11.00.00

REPORT BY: POLICY=BASEPOL

-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---
AVG      790.12 ACTUAL            27.787  SSCHRT  3975  IOC      4233K  CPU 4659.039  CP     542.89
MPL      790.12 EXECUTION         15.761  RESP     2.8  CPU    306468K  SRB 1513.624  AAPCP    0.00
ENDED      9173 QUEUED                1  CONN     1.4  MSO         0  RCT    0.622  IIPCP    2.00
END/S     10.19 R/S AFFIN             0  DISC     1.2  SRB     97415K  IIT   21.116
#SWAPS     6087 INELIGIBLE             0  Q+PEND   0.1  TOT    408116K  HST    0.000  AAP       N/A
EXCTD     15860 CONVERSION             0  IOSQ     0.0  /SEC   453461  AAP      N/A  IIP     97.84
AVG ENC    4.00 STD DEV         8.40.915                             IIP 1308.346
REM ENC    0.00                                         BSRPTN    574
MS ENC     0.00                                         TRX SERV  574
```

SMT active:
MT-1 equivalent service units and service times

SMT active:
Percentage of maximum core capacity used

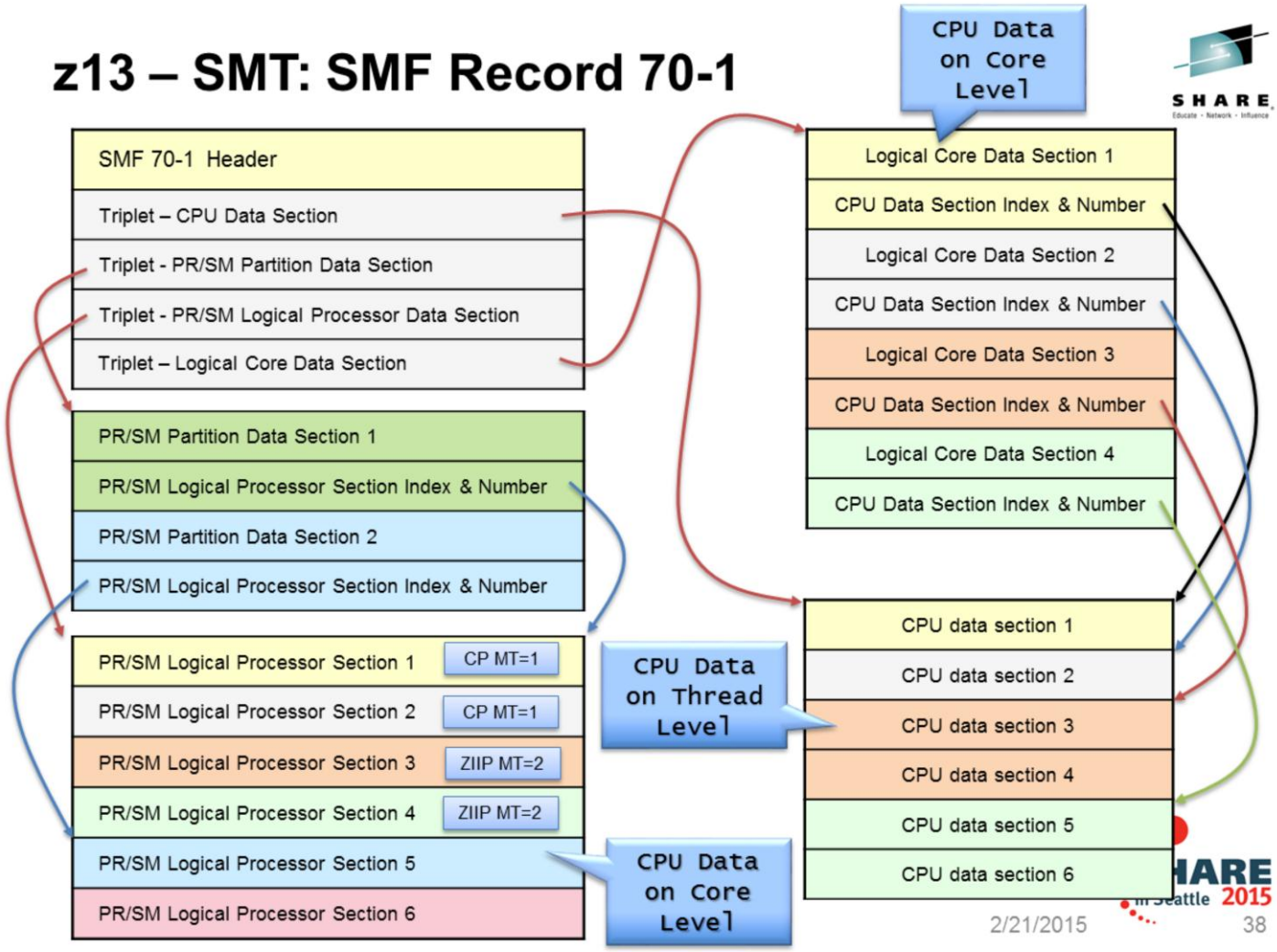
- Pre SMT/ SMT mode inactive:
  - ⇨ Service time : Logical processor CPU time
  - ⇨ APPL%: Percentage of logical processor capacity
- SMT mode active:
  - ⇨ Service time: MT-1 equivalent CPU time
  - ⇨ APPL% : Percentage of maximum core capacity calculated as:

$$APPL\% = \frac{\text{MT}-1 \text{ equivalent CPU time}}{\text{Interval length} \times \text{mCF}} \times 100$$

For MT Mode = 1 ⇨ mCF = 1

- The RMF Postprocessor Workload Activity report (WLMGL) reports the CPU time used by a workload in units of service times and service units. The APPL% metric shows the percentage of logical processor capacity used by the workload.
- With active SMT mode, the service time charged to the workload is based on the MT-1 equivalent CPU time (the CPU time that would have been used in MT-1 mode) . Service units are calculated from MT-1 equivalent CPU time, too.
  The APPL% now represents the percentage of maximum core capacity used by the workload.
- MT-2 APPL% numbers can continue to be used to understand relative core utilization in a given interval, or at times of comparable Maximum Capacity Factors. However, the Maximum Capacity Factor (mCF) needs to be considered when comparing APPL% across different workloads or times with different mCF values.
- If multithreading is disabled for a processor type, an mCF of 1 is used for the APPL% calculation so that the calculation is the same as before introduction of SMT.

z13 – SMT: SMF Record 70-1



- The RMF SMT support enhances the SMF 70-1 record with SMT related fields.
- **CPU Control Data Section:**
  - Processor Class Maximum Capacity Factor Metrics:
    - SMF70MCF (CP), SMF70MCFS (zIIP)
  - Processor Class Capacity Factor Metrics:
    - SMF70CF (CP), SMF70CFS (zIIP)
  - Processor Class Average Thread Density Metric:
    - SMF70ATD (CP), SMF70ATDS (zIIP)
- **PR/SM Partition Data Section:**
  - Maximum Thread Id and MT enabled (SMF70MTID)
- **PR/SM Logical Processor Data Section:**
  - MT inactive → section is on a per logical processor/CPU basis
  - MT active → section is on a per logical core basis
  - MT Idle Time (SMF70MTIT)
- With SMT active, there is no longer a 1:1 mapping between PR/SM Logical Processor data sections and CPU data sections. The PR/SM Logical Processor data sections now represent CPU data on logical Core level, the CPU data sections represent CPU data on logical thread level. To identify the CPU data sections belonging to a logical Core, a new Logical Core data section is introduced.

z13 – SMT: SMF Record 70-1…

SMF record type 70 subtype 1 – Logical Core data section				
Offsets	**Name**	**Len**	**Format**	**Description**
0 0	SMF70_CORE_ID	2	Binary	Core identification.
2 2	SMF70_CORE_FLG	1	Binary	Logical Core Information **Bit Meaning When Set** 0 Core LPAR Busy time is valid. 1-7 Reserved.
3 3		1		Reserved.
4 4	SMF70_CPU_SKIP	2	Binary	The CPU data sections for this core are grouped together in the record. To get to the first CPU data section associated with this logical core, skip over the number of CPU data sections specified by this field, starting at the first CPU data section in the record.
6 6	SMF70_CPU_NUM	2	Binary	Number of CPU data sections for this core. This value represents the number of threads that are active on this core.
8 8	SMF70_PROD	4	Binary	Multithreading core productivity numerator. Divide this value by 1024 to get the multithreading core productivity. A zero value is reported if the core was not configured ONLINE for the complete interval. If SMF70_CPU_NUM is greater than 1, the core productivity represents the percentage of how much work the core resources accomplished while dispatched to physical hardware over the maximum amount of work the core resources could have accomplished while dispatched to physical hardware.
12 C	SMF70_LPAR_BUSY	4	Binary	Multithreading core LPAR Busy Time in milliseconds. This field is valid if bit 0 of SMF70_CORE_FLG is set.

- A new **Logical Core data section** is added to the SMF record 70 subtype 1 when the SMT is active (LOADxx PROCVIEW CORE parameter is in effect).
- The SMF 70 Subtype 1 Individual Header Extension is extended by a new triplet that describes the new Logical Core data sections.
- **Logical Core Data Section Fields:**
- SMF70_CPU_SKIP and SMF70_CPU_NUM can be used to identify the CPU data sections with the thread data of a logical core:
  - Navigate to first CPU on core via a number of CPU data sections to skip (SMF70_CPU_SKIP)
  - Number of CPU data sections for this core (SMF70_CPU_NUM)
- Core Productivity Metric (SMF70_PROD)
- Core LPAR Busy Metric (SMF70_LPAR_BUSY)

z13 – SMT: SMF 70-1 based Overview Conditions

- New qualifier *coreid* is added to support overview reports for core metrics
  - *coreid* is a processor identifier (one or two hexadecimal digits) that either identifies a logical core (when LOADxx PROCVIEW CORE is in effect) or a logical processor (when LOADxx PROCVIEW CPU is in effect).
  - If the qualifier is omitted, the values represent the average of all logical processors/cores

Condition	Name	Qualifier	Source	Algorithm
Percent Multi-Threading core productivity for zIIPs	IIPPROD	coreid	SMF70_PROD	PROD / 1024
Percent Multi-Threading core utilization for zIIPs	IIPUTIL	coreid	SMF70_PROD SMF70PDT SMF70ONT SMF70_LPAR_BUSY	PROD / 1024 multiplied by value of Overview Condition IIPBSY

- For SMF 70-1 based overview reporting there are new overview conditions that can be used to display the MT Core Productivity and MT Core Utilization for a reporting interval.
- The new overview qualifier coreid allows overview reporting at core granularity.

z13 – SMT: SMF 70-1 based Overview Conditions …

- Existing qualifier *cpuid* is changed to support overview reports for logical processor and thread metrics
- *cpuid* is a processor identifier which must be in the format procid[.threadid]
  - procid is a processor identifier (one or two hexadecimal digits) that eitheridentifies a logical core (when LOADxx PROCVIEW CORE is in effect) or a logical processor (when LOADxx PROCVIEW CPU is in effect)
  - threadid is an optional thread identifier (0 or 1) that identifies a thread that is executing on the logical core designated by cpuid. It is ignored when LOADxx PROCVIEW CPU is in effect
  - If LOADxx PROCVIEW CORE is in effect and threadid is omitted, the values represent the average of all threads executing on the logical core
- Examples with PROCVIEW CORE active

  0A specifies logical core id 0A
  3F.0 specifies thread id 0 of logical core id 3F
  A.1 specifies thread id 1 of logical core id 0A
- Examples with PROCVIEW CPU active

  0A specifies logical processor id 0A
  3F.0 specifies logical processor id 3F

- To support overview reporting for logical processor and thread metrics the cupid qualifier can now be specified in the format cpuid[.threadid]:
  - cpuid is a processor identifier (one or two hexadecimal digits) that either identifies a logical core (when LOADxx PROCVIEW CORE is in effect) or a logical processor (when LOADxx PROCVIEW CPU is in effect).
  - threadid is an optional thread identifier (0 or 1) that identifies a thread that is executing on the logical core designated by cpuid. It is ignored when LOADxx PROCVIEW CPU is in effect .
    If LOADxx PROCVIEW CORE is in effect and threadid is omitted, the values represent the average of all threads executing on the logical core.

z13 – SMT: RMF Distributed Data Server Metrics

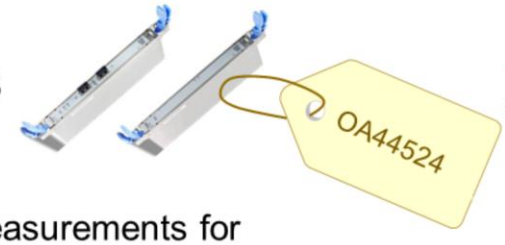
Sysplex

- MVS Image
  - I/O Subsystem
  - Processor
  - Storage
  - Enqueue
  - Operator
  - Subsystems
- CPC
  - LPAR
- Coupling Facility
  - CF Structure

Resource	New metric
Sysplex	► MT mode for IIP by partition ► MT capacity factor for IIP by partition ► MT maximum capacity factor for IIP by partition ► % MT IIP core productivity by partition ► average thread density for IIP by partition ► MT mode for CP by partition ► MT capacity factor for CP by partition ► MT maximum capacity factor for CP by partition ► % MT CP core productivity by partition ► average thread density for CP by partition
LPAR	► MT Mode IIP ► MT capacity factor for IIP ► MT maximum capacity factor for IIP ► % MT IIP core productivity ► average thread density for IIP ► MT Mode CP ► MT capacity factor for CP ► MT maximum capacity factor for CP ► % MT CP core productivity ► average thread density for CP

- The RMF Distributed Data Server (DDS) provides a new set of MT metrics for each processor class:
  - MT Mode
  - MT Core Productivity
  - MT Maximum Capacity Factor
  - MT Capacity Factor
  - Average Thread Density
- These metrics are available as single valued metrics for the LPAR resource and as list valued metrics for the SYSPLEX resource.

z13 - RMF PCIE Enhancements



- z13 introduces new PCIE performance measurements for RDMA-over-converged-ethernet (RoCE Express) and zEnterprise data compression (zEDC) devices
  - Existing DMA read/write measurements are replaced by new PCIE function type specific measurements
  - For RoCE Express devices, there are four new measurements
    - **Received-Bytes** — No. of bytes received on the external ethernet interface
    - **Transmitted-Bytes** — No. of bytes transmitted on the external ethernet interface
    - **Received-Packets** — No. of packets received on the external ethernet interface
    - **Transmitted-Packets** — No. of packets transmitted on the external ethernet interface
  - For zEDC devices, there are two new measurements
    - **Consumed-Work-Units** — No. of work units processed by the PCI function
    - **Maximum Work Units** — Maximum no. of work units that the PCI function is capable of processing per second
- With zEC12 / zBC12, the existing DMA Read/Write metrics are still maintained
    - **DMA Reads** — No. bytes transferred from DMA address spaces to PCIE function
    - **DMA Writes** — No. bytes transferred from PCIE function to DMA address spaces

- RMF PCIE monitoring and reporting functionality is enhanced to support new measurements for PCIe-attached RoCE and zEDC devices configured on z13.

z13 - Support of Crypto Express5 Card

OA43493

```
                                    C R Y P T O   H A R D W A R E   A C T I V I T Y
z/OS V2R2                 SYSTEM ID TRX2              DATE 02/01/2014            INTERVAL 14.59.998
                                  RPT VERSION V2R2 RMF       TIME 11.00.00              CY
-------- CRYPTOGRAPHIC CCA COPROCESSOR ----
           -------- TOTAL --------  KEY-GEN
TYPE   ID   RATE  EXEC TIME  UTIL%    RATE
CEX5C   0   2.15    227.8    48.9     2.15

-------- CRYPTOGRAPHIC PKCS11 COPROCESSOR ----------------------------------------------------
           -------- TOTAL --------    ----------- OPERATIONS DETAILS ------------
TYPE   ID   RATE  EXEC TIME  UTIL%    FUNCTION          RATE     EXEC TIME     UTIL%
CEX5P   1  446.5    0.243     8.3     ASYM FAST        274.3        0.175       2.4
                                      ASYM GEN          0.00        0.000       0.0
                                      ASYM SLOW        120.3        0.405       5.3
                                      SYMM COMPLETE     0.00        0.000       0.0
                                      SYMM PARTIAL     51.89        0.398       0.6

-------- CRYPTOGRAPHIC ACCELERATOR ------------------------------------------------------------------------------
           -------- TOTAL -------            -- ME-FORMAT RSA OPERATIONS --  -- CRT-FORMAT RSA OPERATIONS --
TYPE   ID   RATE  EXEC TIME UTIL%    KEY      RATE     EXEC TIME     UTIL%      RATE      EXEC TIME     UTIL%
CEX5A   2 1335.5    0.151    30.3    1024     678.2     0.225        14.2     544.4        0.145        5.8
                                     2048      0.00     0.000         0.0      22.6        0.465        4.8
                                     4096      0.00     0.000         0.0      90.3        0.378        5.5

-------- ICSF SERVICES ------------------------------------------------------------------------------------------
        ---- ENCRYPTION -----     ---- DECRYPTION -----     -------- HASH ---------     ------ PIN -------
         SDES     TDES     AES     SDES     TDES     AES     SHA-1  SHA-256  SHA-512    TRANSLATE   VERIFY
RATE     0.00     0.00    0.00     0.00     0.00    0.00      0.00     0.00     0.00         0.00      0.00
SIZE     0.00     0.00    0.00     0.00     0.00    0.00      0.00     0.00     0.00

        ----- MAC ------   ---- AES MAC ---    --- RSA DSIG ---    --- ECC DSIG ---    - FORMAT PRESERVING ENCRYPTION -
        GENERATE  VERIFY   GENERATE  VERIFY    GENERATE  VERIFY    GENERATE  VERIFY    ENCIPHER   DECIPHER    TRANSLATE
RATE        0.00     0.00       0.00     0.00       0.00     0.00       0.00     0.00       0.00       0.00         0.00
SIZE        0.00     0.00       0.00     0.00                                           0.00       0.00         0.00
```

Crypto Express5 Card reported with same set of measurements as Crypto Express4

Additional ICSF service statistics

- RMF new function APAR OA43493 introduces the RMF support of the Crypto Express5S feature on the z13
- In detail, RMF collects and reports performance measurements for operations executed on Crypto Express5S CCA coprocessors (CEX5C) , PKCS11 coprocessors (CEX5P) and accelerators (CEX5A).
- The crypto measurements are stored to SMF 70 subtype 2 data sections:
  - CEX5C measurements are stored in the Cryptographic Coprocessor data section,
  - CEX5P measurements are stored in the Cryptographic PKCS11 Coprocessor data section and
  - CEX5A measurements are stored in the Cryptographic Accelerator data section.
- The Postprocessor Crypto Activity report provides the crypto measurements from the SMF 70 subtype 2 data sections in the corresponding report sections.
- The ICSF SERVICES report section displays request rates for the new ICSF activities:
  - RSA Digital Signature Generate callable services
  - RSA Digital Signature Verify callable services
  - ECC Digital Signature Generate callable services
  - ECC Digital Signature Verify callable services
  - AES MAC Generate callable servies
  - AES MAC Verify callable servies
  - FPE Encipher callable services
  - FPE Decipher callable services
  - FPE Translate callable services
- With new function APAR OA43493 the RMF support is available for z/OS V1.13 and z/OSV2.1.

RMF Enhancements at a Glance

- IBM zEnterprise EC12 Support
  - Statistics for Flash Memory and Pageable Large Pages
  - Support of Coupling Facility Flash Memory
  - z Enterprise Data Compression Express Reporting (zEDC)
  - Shared Memory Communication Reporting (SMC-R)
  - Support of I/O Interrupt Delay Time Facility
  - Extended Infiniband Link Reporting
  - Support of Crypto Express4 Card

- IBM z13 Support
  - RMF support for SMT, Crypto Express5S …
  - PCIe reporting enhancements

- Exploitation of System z Integrated Information Processors
  - zIIP Usage Option for Monitor III Gatherer

- Postprocessor XML formatted Reports
  - Transition to XML Format almost complete
  - Advanced Sorting and Filtering Capabilities

- RMF XP Enhancements
  - SMF Type 104 Recording for all Platforms

- z/OSMF Resource Monitoring
  - Context sensitive Application Linkage to WLM
  - Historical Reporting & Spreadsheet Export

RMF Monitor III zIIP Exploitation

- With z/OS V2R1 RMF, the Monitor III Data Gatherer (RMFGAT) can partially offload work to zIIP processors
- By default the RMF Monitor III Data Gatherer (RMFGAT) is enabled for zIIP exploitation
- When at least one zIIP processor is online for an LPAR, RMFGAT is partially offloading work to this processor without any further user interaction
- The RMFGAT zIIP exploitation can be controlled initially by means of the new Monitor III parmlib option ZIIPUSE

```
SYNC(00)                /* MINTIME SYNCHRONIZATION              */
SYSOUT(A)               /* MESSAGES TO SYSOUT CLASS A           */
WSTOR(32)               /* SIZE OF INSTORAGE BUFFER (IN MB)     */
ZIIPUSE                 /* PARTIAL USE OF zIIP ENGINES          */
IOSUB                   /* I/O SUBSYSTEM GATHERING ACTIVE       */
CFDETAIL                /* COUPLING FACILITY DETAILS            */
CACHE                   /* ACTIVATE CACHE GATHERING             */
VSAMRLS                 /* ACTIVATE VSAM RLS GATHERING          */
OPD                     /* ACTIVATE OMVS PROCESS DATA GATHERING */
```

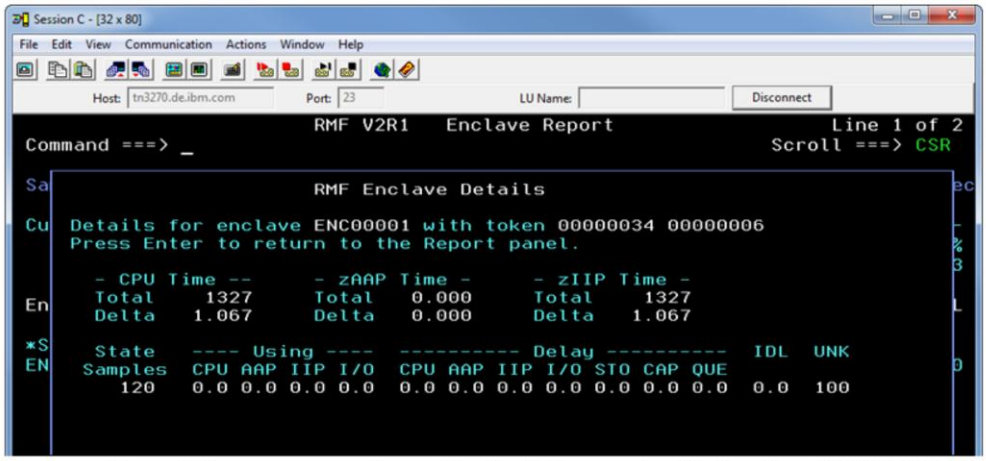
New Option ZIIPUSE

- The RMFGAT zIIP exploitation can be activated/deactivated dynamically by means of the following command: F RMF,F III,ZIIPUSE/NOZIIPUSE

- The RMF Monitor III data gatherer (RMFGAT) has been entitled to partially run on IBM System z Integrated Information Processors (zIIPs).
- A new Monitor III gatherer option ZIIPUSE/NOZIIPUSE is introduced to determine whether the RMFGAT address space is eligible for zIIP exploitation.

RMF Monitor III zIIP Exploitation...

- When at least one zIIP processor is recognized by the Monitor III gatherer, RMFGAT will schedule an Enclave SRB and offloads the Coupling Facility gathering to the zIIP processor



- In case the zIIP processor is activated dynamically by the CONFIG CPU(xx),ONLINE command, RMFGAT can exploit this processor starting with the next MINTIME
- Installations without Coupling Facilities (e.g. Monoplex) won't see significant RMFGAT zIIP activity

- When at least one zIIP processor is recognized by the Monitor III gatherer, RMFGAT will schedule an Enclave SRB and offloads the RMF Coupling Facility data gathering to the zIIP processor.
- The Enclave will be displayed in RMF Monitor III Enclave report. In the Enclave Details pop-up panel you will see the RMFGAT CPU time offloaded to zIIP.
- If the zIIP will be taken online dynamically, it will be recognized by RMFGAT and the zIIP will be exploited with the start of the next Monitor III Mintime.
- When no zIIPs are online, RMFGAT will neither create an enclave nor schedule any zIIP eligible SRB work. That means, no zIIP time on CP will appear in any performance reports.
- Since RMFGAT only offloads the Coupling Facility data gathering, installations without Coupling Facilities will not benefit from the zIIP exploitation.

RMF Enhancements at a Glance

- IBM zEnterprise EC12 Support
  - Statistics for Flash Memory and Pageable Large Pages
  - Support of Coupling Facility Flash Memory
  - z Enterprise Data Compression Express Reporting (zEDC)
  - Shared Memory Communication Reporting (SMC-R)
  - Support of I/O Interrupt Delay Time Facility
  - Extended Infiniband Link Reporting
  - Support of Crypto Express4 Card

- IBM z13 Support
  - RMF support for SMT, Crypto Express5S …
  - PCIe reporting enhancements

- Exploitation of System z Integrated Information Processors
  - zIIP Usage Option for Monitor III Gatherer

- Postprocessor XML formatted Reports
  - Transition to XML Format almost complete
  - Advanced Sorting and Filtering Capabilities

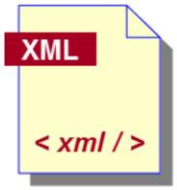
- RMF XP Enhancements
  - SMF Type 104 Recording for all Platforms

- z/OSMF Resource Monitoring
  - Context sensitive Application Linkage to WLM
  - Historical Reporting & Spreadsheet Export

Postprocessor XML Formatted Reports

Rationale:

- RMF Postprocessor reports are limited to a page width of 132 characters
- No state-of-the-art display capability of Postprocessor reports
- No easy access to RMF Postprocessor data for application programs
  - cumbersome to parse the text output
  - each report has its own layout



RMF Postprocessor reports can now be generated in XML Format

- In the past, users of the RMF Postprocessor had to struggle with a couple of shortcomings:
- The standard text output was limited to to a width of 132 characters. Therefore some reports appeared somehow squeezed. This applies in particular to Overview Reports with a high number of OVW conditions
- The basic text format didn't exploit todays presentation capabilties like resizing windows, scrolling back and forth etc.
- Depending on the report type, the report layouts were quite different. Hence, API programs need to supply specific logic in order to extract metrics out of the reports
- Starting with z/OS V1R11 RMF, the Postprocessor is able to generate a subset of the reports alternatively
- in XML format.

Postprocessor XML Formatted Reports...

z/OS V1R11 RMF	z/OS V1R12 RMF	z/OS V1R13 RMF	z/OS V2R1 RMF
CPU Activity CRYPTO Activity FICON Director Activity ESS Disk Systems Activity OMVS Kernel Activity report OVERVIEW Report	DEVICE Activity WORKLOAD Activity	PAGING Activity SDELAY (XML only)	CACHE Subsystem Activity CF Activity CHANNEL Path Activity ENQUEUE Avtivity HFS Statistics IOQ Activity PCIE Activity (XML-only) PAGESP Activity SDEVICE Activity report VSTOR Activity XCF Activity

- Summary and Exception reports as well as interval reports based on data collected by a Monitor II background session are not available in XML format
- The XML format is the preferred RMF Postprocessor Report format for the future
- The XML Format supersedes the Text format. New Reports might not be implemented in Text format

- It was RMF's objective to introduce the XML format for the majority of Postprocessor reports during the past releases. This work is now completed with RMF for z/OS V2.1.
- With the first stage in RMF for z/OS 1.11, those reports have been selected where no similar Monitor III report is available – since Monitor III reports are already XML enabled by means of the Data Portal. These are the following Postprocessor single-system reports:
  - CPU Activity report
  - CRYPTO Hardware Activity report
  - FICON Director Activity report
  - OMVS Kernel Activity report,
  - ESS Disk Systems Activity report
- In addition, Postprocessor Overview reports belong also to the first stage since multiple Overview conditions have caused nasty line-feeds in the old text format.
- RMF for z/OS 1.12 extends the list of Postprocessor XML reports with another single- system report and the first sysplex report:
  - DEVICE Activity report (single-system report)
  - WLMGL Workload Activity report (sysplex report)
- z/OS 1.13 added the PAGING Activity report and the Serialization Delay report (SDELAY) which was the first report that is only available in XML format. Now with z/OS 2.1, the list of available reports is completed with:
  - CACHE Subsystem Activity report
  - CF Coupling Facility Activity report
  - CHANnel Path Activity report
  - ENQueue Activity report
  - HFS Hierarchical File System report
  - IOQ I/O Queuing Activity report
  - PCIE Activity Report
  - PAGESP Page Data Set Activity report
  - SDEVICE Shared Device Activity report
  - VSTOR Virtual Storage Activity report
  - XCF Cross-System Coupling Activity report

Postprocessor XML Formatted Reports...

- The generation of Postprocessor reports in XML format is controlled by the new ddnames XPRPTS, XPXSRPTS and XPOVWRPT
- If the XML output is routed to permanent data sets rather than to SYSOUT, define the data set with RECFM=VB and LRECL between 256 and 8192. Specify an appropriate BLKSIZE.

ddname	Contents	Allocations	Notes
XPRPTS	Combined **single-system** report in XML format	One ddname for one data set to contain all single system reports for each interval during the session.	There is no dynamic allocation of this ddname, you have to define it explicitly if you want to get all reports in XML format into one data set or output class. If you define this ddname, no MFRnnnnn files are created. If you define this ddname and PPRPTS, no XML output in file XPRPTS is created.
XPOVWRPT	Combined **Overview** report in XML format	One ddname for one data set to contain all overview reports for each system included in the input data.	There is no dynamic allocation of this ddname, you have to define it explicitly if you want to get all overview reports in XML format into one data set or output class. If you define this ddname, no PPORPnnn files are created.
XPXSRPTS	Combined **sysplex-wide** report in XML format	One ddname for one data set to contain all sysplex reports for each interval included in the input data.	There is no dynamic allocation of this ddname, you have to define it explicitly if you want to get all reports in XML format into one data set or output class. If you define this ddname, no MFRnnnnn files are created. If you define this ddname and PPXSRPTS, no XML output in file XPXSRPTS is created.

- The request for XML output from the RMF Postprocessor is triggered by means of the existence of a DD card for XPRPTS (for single-system reports) , XPOVWRPT (for Overview reports) and/or XPXSRPTS (for sysplex-wide reports) in the JCL for the Postprocessor job.
- If one of these DD statements are defined in the JCL, then XML output will be generated.
- The following rules apply, if requests for text and XML are given simultaneously:
  - If the user specifies a DD card for cumulative text output, i.e. PPRPTS, as well as XPRPTS, then only text output will be generated into PPRPTS and no XML is produced.
  - If the user specifies the PPXSRPTS ddname for cumulative text output together with XPXSRPTS, no XML output is generated.
  - If single-system reports are requested on the REPORTS control statement which are not XML enabled and the DD card XPRPTS is defined, no output is generated for them.

HTTP API to access Postprocessor XML Reports

- Web browser can be used as Postprocessor Data Portal
- All RMF Postprocessor XML formatted reports supported
- Application programs can use Distributed Data Server (DDS) HTTP API to retrieve RMF Postprocessor XML reports

http://boesysf:8803/gpm/rmfpp.xml?reports="CPU"&date=20100702,20100702 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

RMF Postprocessor Interval Report [SYSD] : CPU Activity Report

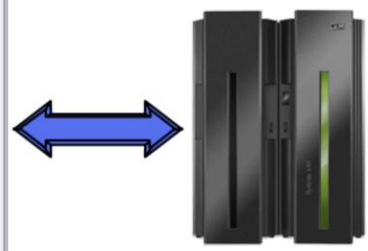
z/OS V1R10 Start: 07/02/2010-11.45.00 Interval: 14:59:999 minutes

V1R12 RMF End: 07/02/2010-12.00.00 Cycle: 1000 milliseconds

CPU ACTIVITY

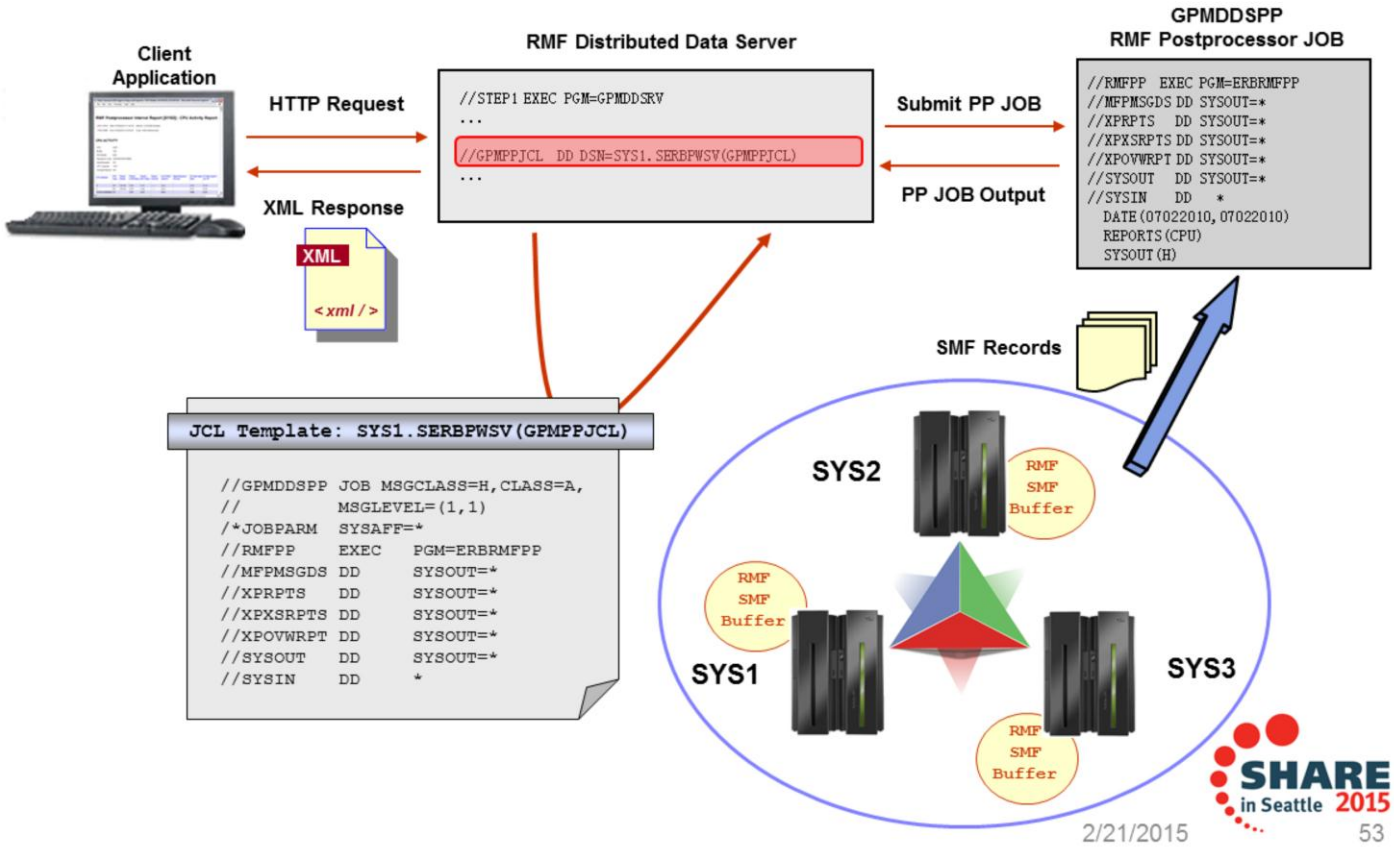
CPU	2097
Model	722
H/W Model	E40
Sequence Code	0000000000070B82
HiperDispatch	NO
CPC Capacity	1631
Change Reason	N/A

CPU Number	CPU Type	Time% Online	Time% LPAR Busy	Time% MVS Busy	Time% Parked	LOG PROC Share%	HiperDispatch Priority	I/O Interrupts Rate	I/O Interrupts% via TPI
0	CP	100.00	3.50	3.47	------	32.2		4.61	0.19
1	CP	100.00	3.52	3.48	------	32.2		5.07	0.46
TOTAL/AVERAGE	CP		3.51	3.47		64.4		9.68	0.33



- In most installations, the access to historical data is needed for in depth performance analysis. This allows to keep track whether a critical situation has been persistent or not.
- The existing HTTP API of the RMF Distributed Data Server (DDS) already provides a sysplex-wide access of the data collected by RMF Monitor III. With RMF for z/OS 1.12 this API is extended by the capability to grant instant and easy access to RMF long-term historical data as reported by the RMF Postprocessor.
- Application programs can exploit the extended DDS HTTP API by sending standard URL requests for historical RMF Postprocessor data to the DDS.
- Since DDS returns the requested data as XML document, a web browser can act as an RMF Postprocessor Data Portal to access historical RMF data which is generated by the RMF Postprocessor in XML format.

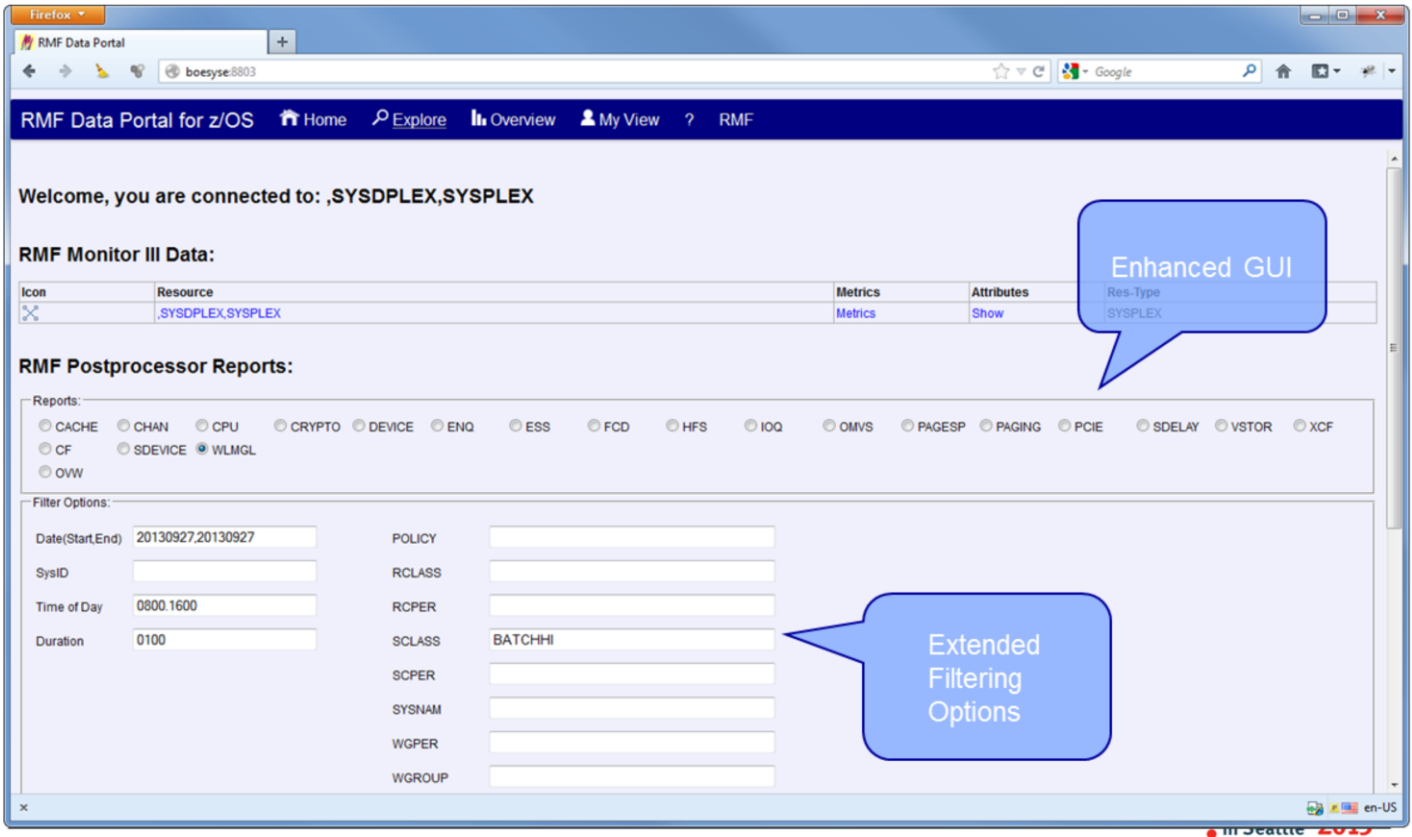
HTTP API to access Historical Data...



- To get access to RMF Postprocessor data provided by the DDS, the GPMSERVE started task is extended with an additional DD card:
  //GPMPPJCL DD DISP=SHR,DSN=SYS1.SERBPWSV(GPMPPJCL)
- This DD statement points to a JCL template which contains all JCL statements necessary to start a Postprocessor job from the RMF Distributed Data Server. The JCL template is stored in SYS1.SERBPWSV(GPMPPJCL).
- You must adapt or replace the GPMPPJCL member to suit your installation, ensuring that the DDS is able to run RMF Postprocessor jobs. If you do not want to request Postprocessor data with the DDS, you can omit the GPMPPJCL ddname.
- The dataset or the member specified in the GPMPPJCL DD card can be replaced with private ones containing the JCL template for running RMF Postprocessor from the DDS.
- The JCL template that is used by DDS to start the Postprocessor JOB does not contain any Postprocessor control statements. These will be dynamically created by DDS depending on the HTTP request parameters.
- For a detailed description of the DDS HTTP API please refer to the z/OS RMF Programmer's Guide.

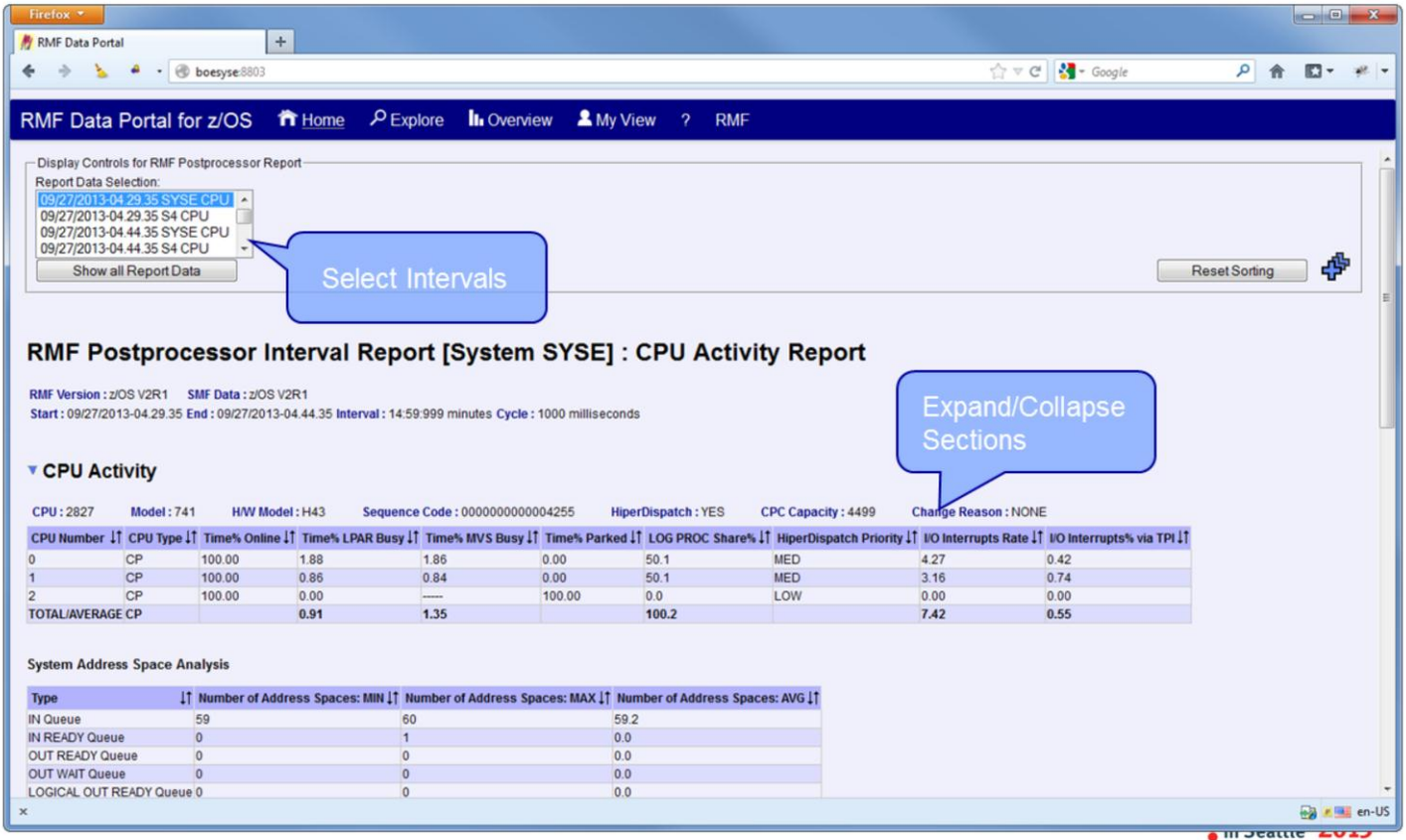
- Notes:
- The Postprocessor API functionality is only available with JES2 installed.
- The DD card specifications for the XML output datasets must not be changed. To be able to retrieve the Postprocessor output, DDS needs these SYSOUT data sets to reside on spool.

Postprocessor XML Formatted Reports...



- The RMF Data Portal for z/OS allows to generate and display and RMF Postprocessor XML report:
- When you click the Explore button on the welcome screen, the Data Portal displays a list of Postprocessor XML enabled reports.
- The selection of a Postprocessor report triggers an HTTP request that will retrieve, format and display the Postprocssor report in the web browser
- With RMF for z/OS 2.1 the GUI of the RMF Data Portal for z/OS was enhanced:
- For each Postprocessor report type you can now specify additional options that will be used as Postprocessor report control statements.
- These options allow to filter the generated Postprocessor XML report, so that the amount of XML output can be reduced.
- This helps to improve the performance to generate and display a Postprocessor XML report in the RMF Data Portal for z/OS.

Postprocessor XML Formatted Reports...



- The new stylesheets and JAVA script code, used to format and display the Postprocssor XML report in the
- web browser introduces new functionality that helps to analyse the RMF performance data:
- You can now select a single report interval or a set of report intervals to focus on the RMF performance data for specific report interval(s).
- You can expand/collapse report section(s) so that it‘s easier to get an overview about the reported data.

Postprocessor XML Formatted Reports...

The RMF Data Portal is currently limited to the SMF Buffer of the RMF Sysplex Data Server. No SMF Dump Data Sets can be specified

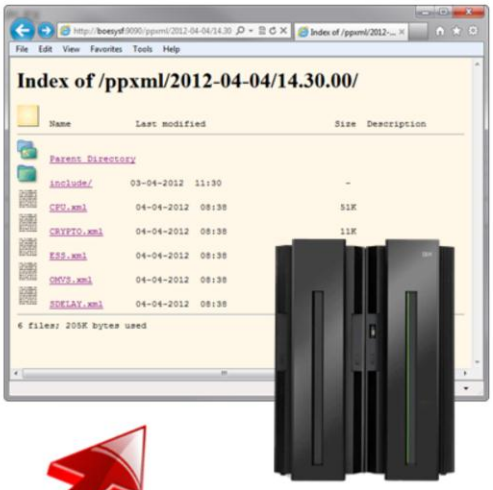
```
//SHARAnnX JOB (DE03141,,),'SHARAnn',CLASS=A,USER=SHARAnn,
//        MSGCLASS=H,MSGLEVEL=(1,1),NOTIFY=SHARAnn
//*
//XMLPP    PROC REPORT=
//RMFPP    EXEC PGM=ERBRMFPP,COND=(4,LT,GETSMF)
//MFPINPUT DD   DISP=(OLD,PASS),DSN=*.RMFSORT.SORTOUT
//MFPMSGDS DD   SYSOUT=*
//XPRPTS   DD   PATH='/sharelab/sharann/xmlpp/&REPORT..xml',
//         PATHOPTS=(OWRONLY,OCREAT,OTRUNC),
//         PATHMODE=(SIRUSR,SIWUSR,SIRGRP),FILEDATA=TEXT
//XPXSRPTS DD   PATH='/sharelab/sharann/xmlpp/&REPORT..xml',
//         PATHOPTS=(OWRONLY,OCREAT,OTRUNC),
//         PATHMODE=(SIRUSR,SIWUSR,SIRGRP),FILEDATA=TEXT
//         PEND
//PPCPU    EXEC PROC=XMLPP,REPORT=CPU
//RMFPP.SYSIN DD   *
  REPORTS(CPU)
//PPWLMGL  EXEC PROC=XMLPP,REPORT=WLMGL
//RMFPP.SYSIN DD   *
  SYSRPTS(WLMGL(SCPER,POLICY))
//
```

Execute the Postprocessor JCL and redirect the output to any HFS directory

- The RMF Postprocessor report generation via the RMF Data Portal for z/OS is currently limited to the SMF records in the SMF Buffer of the RMF Sysplex Data Server since no SMF Dump Data Sets can be specified in the HTTP request for the Postprocessor API.
- However, you can use an RMF Postprocessor JOB to generate an XML report from any SMF source and redirect the XML output to a permanent file in the USS filesystem.
- Now that you have generated a report, there are multiple ways to view the XML output:
  - Transfer the XML file in ASCII mode to your windows workstation and use the RMF Postprocessor XML Toolkit to display the report in your web browser as an HTML document.
  - Or, if you don‘t want to transfer the XML report to your workstation, you can think about using an SMB Server to map the USS directory tree to a network drive or you can setup an HTTP server that runs on z/OS and allows to view the XML reports in the USS directory.

Postprocessor XML Report Access via IBM HTTP Server

RMF Postprocessor Reports:
Everywhere and Anytime



- ✓ Instant Access to all Reports
- ✓ No Maintenance of raw Data Archives

- Since the reports are pure XML, an arbitrary web browser is perfectly suited to display the reports in a state-of-the art fashion.
- Once you have created a USS directory with XML reports by means of a Postprocessor job, you can use the IBM HTTP Server instantly as gateway.
- The reports can remain in the USS directory of your z/OS system and you can forget about cumbersome downloads.
- Once the HTTP server setup is done by means of a few entries in the servers configuration file (httpd.conf), you can easily navigate within the USS directory structure.
- If you select the report of your choice, it will be displayed instantly in the browser window.

Creating RMF Postprocessor XML Reports

Set up the IBM HTTP Server for Remote RMF Report Access

Friday, March 6, 2015

Session 16799

RMF Enhancements at a Glance

- IBM zEnterprise EC12 Support
  - Statistics for Flash Memory and Pageable Large Pages
  - Support of Coupling Facility Flash Memory
  - z Enterprise Data Compression Express Reporting (zEDC)
  - Shared Memory Communication Reporting (SMC-R)
  - Support of I/O Interrupt Delay Time Facility
  - Extended Infiniband Link Reporting
  - Support of Crypto Express4 Card
- IBM z13 Support
  - RMF support for SMT, Crypto Express5S …
  - PCIe reporting enhancements
- Exploitation of System z Integrated Information Processors
  - zIIP Usage Option for Monitor III Gatherer
- Postprocessor XML formatted Reports
  - Transition to XML Format almost complete
  - Advanced Sorting and Filtering Capabilities
- RMF XP Enhancements
  - SMF Type 104 Recording for all Platforms
- z/OSMF Resource Monitoring
  - Context sensitive Application Linkage to WLM
  - Historical Reporting & Spreadsheet Export

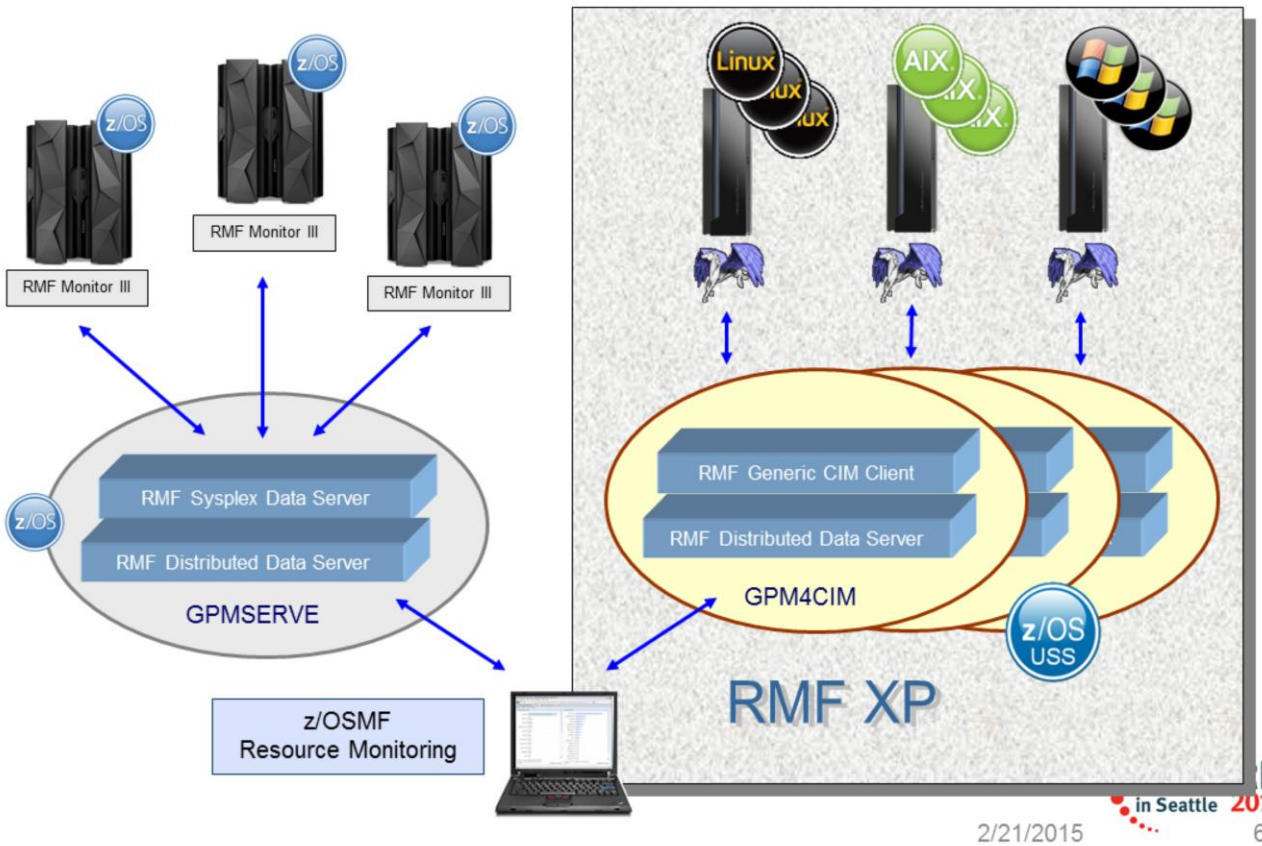
RMF XP Enhancements

- RMF XP is the solution for Cross Platform Performance Monitoring
- RMF XP supports the Operating Systems running on
  - x Blades
  - p Blades

- In addition RMF XP supports Linux on System z
  - LPAR Mode
  - VM Guest Mode

- RMF XP is the new solution to monitor the performance of heterogeneous environments. RMF XP supports the operating systems running on the IBM zEnterprise Bladecenter Extension:
  - AIX on System p
  - Linux on System x
  - Windows on System x (New with z/OS 2.1)
- In addition, Linux on System z is supported as well

RMF XP – Component Overview



- The core component of RMF XP is the GPM4CIM started task.
- Similar to the existing Distributed Data Server for z/OS (aka GPMSERVE) the GPM4CIM started task can receive HTTP requests and sends back responses in terms of structured XML documents.
- Since the GPM4CIM started task runs in the z/OS Unix System Services (USS) environment, at least one z/OS system is needed to utilize the RMF XP component.

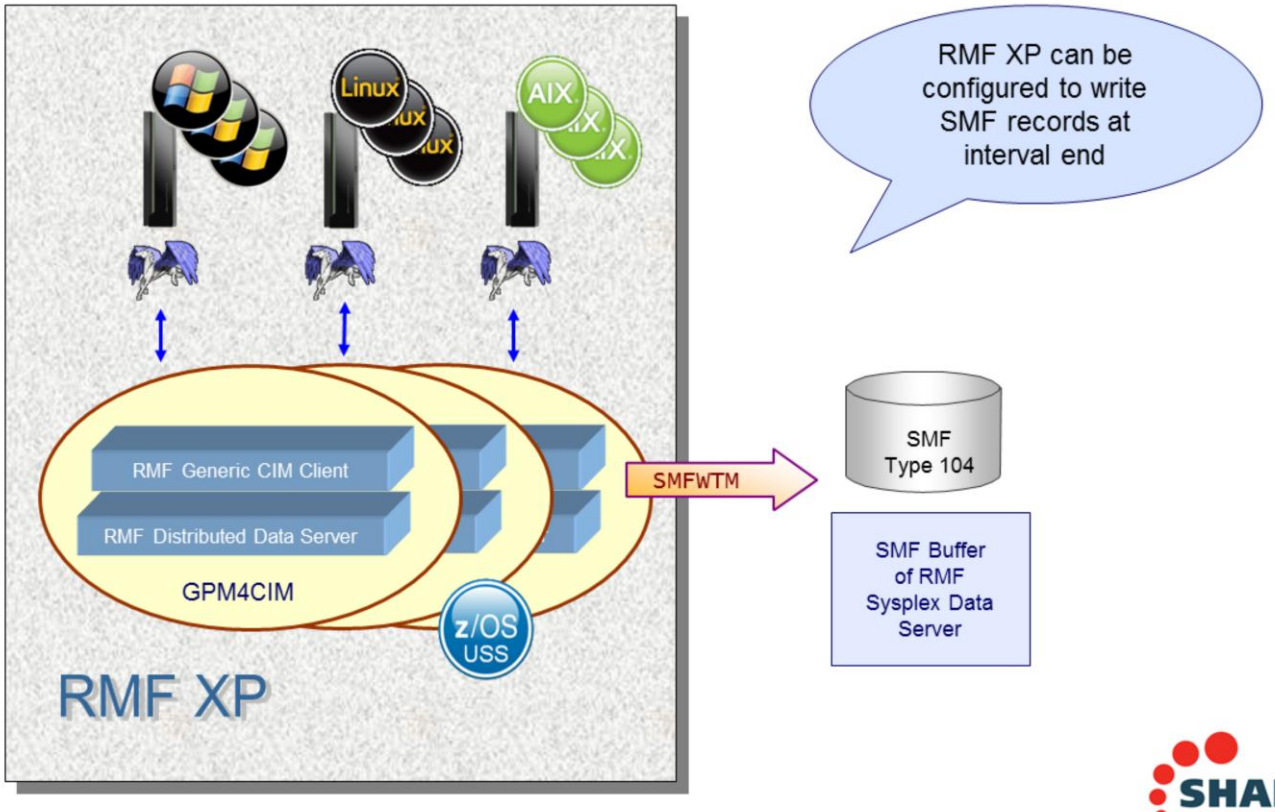
RMF XP – SMF Recording Facility

Rationale

- Data source for after the facts analysis and accounting
- SMF records are the reliable standard on z/OS for decades
- One consitent repository for z/OS and distributed platforms
- Manage z/OS AND distributed platforms from z/OS
- Well-proven SMF postprocessing tools are already in place
- RMF infrastructure can be reused in terms of the RMF Sysplex Dataserver

- RMF Cross Platform Monitoring (RMF XP), introduced with z/OS V1R13 RMF, provides an integrated performance monitoring solution for heterogeneous operating systems running on IBM system z or zBX hardware
- The GPM4CIM server collects the data from the remote AIX, Linux and Windows systems by means of the standard CIM client API.
- However, beyond of the online monitoring capabilities for the new platforms, there was no possibility to store the data persistently for later reporting.
- Especially for z/OS customers, SMF records are the standard repository for all kinds of long term measurement and event data.
- With z/OS V2R1 the RMF GPM4CIM component is extended with the new SMF persistence capability as an additional vehicle for long-term performance analysis and capacity planning.

RMF XP & SMF Records



- Performance data collected from the distributed operating systems can be written to the new SMF Record Type 104 (x'68').
- The new SMF Record Type 104 is written to the currently active SMF Dataset of the z/OS system running the GPM4CIM started task.
- Alternatively a copy of the SMF record image can be written to the SMF Buffer of the RMF Sysplex Data Server. This alternative is not an RMF XP capability, it just exploits existing RMF infrastructure.

RMF XP & SMF Records…

One Subtype per Metric Category

- AIX_SYSTEM_COMPLEX
  - AIX_IMAGE
    - ACTIVE_MEMORY_EXPANSION
    - ACTIVE_MEMORY_SHARING
    - ALL_DISKS
      - DISK
    - ALL_LOGICAL_PROCESSORS
      - LOGICAL_PROCESSOR
    - ALL_NETWORK_PORTS
      - NETWORK_PORT
    - ALL_LOCAL_FILE_SYSTEMS
      - LOCAL_FILE_SYSTEM
    - ALL_PROCESSES
      - PROCESS
    - ALL_SHARED_ETHERNET_ADAPTERS
      - SHARED_ETHERNET_ADAPTER
    - ALL_VIRTUAL_TARGET_DEVICES
      - VIRTUAL_TARGET_DEVICE
    - PARTITION
    - MEMORY

Subtypes 1-12

- XLINUX_SYSTEM_COMPLEX
  - XLINUX_IMAGE
    - ALL_DISKS
      - DISK
    - ALL_LOCAL_FILE_SYSTEMS
      - LOCAL_FILE_SYSTEM
    - ALL_IP_PROTOCOL_ENDPOINTS
      - IP_PROTOCOL_ENDPOINT
    - ALL_LOGICAL_PROCESSORS
      - LOGICAL_PROCESSOR
    - ALL_NETWORK_PORTS
      - NETWORK_PORT
    - ALL_PROCESSES
      - PROCESS
    - ALL_KVM_GUESTS
      - KVM_GUEST
    - ALL_XEN_GUESTS
      - XEN_GUEST

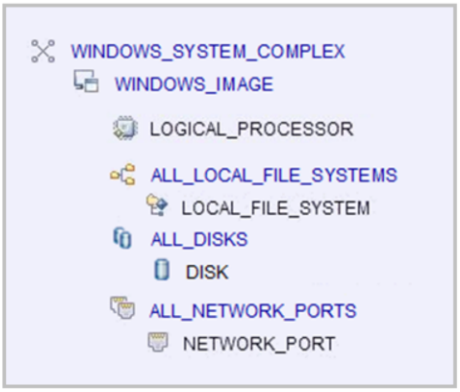
Subtypes 20-31

- ZLINUX_SYSTEM_COMPLEX
  - CEC
    - LPAR
  - ZLINUX_IMAGE
    - ALL_DISKS
      - DISK
    - ALL_LOCAL_FILE_SYSTEMS
      - LOCAL_FILE_SYSTEM
    - ALL_IP_PROTOCOL_ENDPOINTS
      - IP_PROTOCOL_ENDPOINT
    - ALL_LOGICAL_PROCESSORS
      - LOGICAL_PROCESSOR
    - ALL_NETWORK_PORTS
      - NETWORK_PORT
    - ALL_PROCESSES
      - PROCESS
    - ALL_CHANNELS
      - CHANNEL
    - ALL_VOLUMES
      - VOLUME

Subtypes 40-53

- The type 104 SMF record – RMF Distributed Platform Performance Data – serves as a container for all measurement data of non z/OS platforms.
- For each platform, the system writes the following range of subtypes:
  - Subtype 1-12 AIX on System p performance data
  - Subtype 20-31 Linux on System x performance data
  - Subtype 40-53: Linux on System z performance data
  - Subtype 60-64: Windows on System x performance data

RMF XP & SMF Records…



Subtypes 60-64

One Subtype per Metric Category

RMF XP & SMF Records…

One Subtype per Metric Category

AIX on System p	ST	Linux on System x	ST	Linux on System z	ST
AIX_ActiveMemoryExpansion	1	Linux_IPProtocolEndpoint	20	Linux_IPProtocolEndpoint	40
AIX_Processor	2	Linux_LocalFileSystem	21	Linux_LocalFileSystem	41
AIX_ComputerSystem	3	Linux_NetworkPort	22	Linux_NetworkPort	42
AIX_Disk	4	Linux_OperatingSystem	23	Linux_OperatingSystem	43
AIX_NetworkPort	5	Linux_Processor	24	Linux_Processor	44
AIX_FileSystem	6	Linux_UnixProcess	25	Linux_UnixProcess	45
AIX_Memory	7	Linux_Storage	26	Linux_Storage	46
AIX_OperatingSystem	8	Linux_KVM	30	Linux_zCEC	50
AIX_Process	9	Linux_Xen	31	Linux_zLPAR	51
AIX_SharedEthernetAdapter	10			Linux_zChannel	52
AIX_ActiveMemorySharing	11			Linux_zECKD	53
AIX_VirtualTargetDevice	12				

- One specific subtype is used to collect one individual CIM metric category according to the CIM data model on the affected platform.
- The CIM metric category, in return, is mapped to the resource models used by RMF XP on the supported platforms. For example, metrics from the CIM metric category AIX_ActiveMemoryExpansion are related to metrics from the ACTIVE_MEMORY_EXPANSION resource used by RMF XP for AIX on System p.
- You will find a complete mapping of CIM Metric Categories to RMF XP Resource Types in the description of the SMF 104 record in the ‚z/OS V2R1.0 MVS System Management Facilities (SMF)‘ book.

RMF XP & SMF Records…

Windows on System x	ST
Windows_LocalFileSystem	60
Windows_NetworkPort	61
Windows_OperatingSystem	62
Windows_Processor	63
Windows_Storage	64

One Subtype per Metric Category

RMF XP & SMF Records…

SMF Header with common Layout

Offsets		Name	Length	Format	Description
Common header for SMF record type 104					
0	0	SMF104LEN	2	binary	Record length. This field and the next field (total of four bytes) form the RDW (record descriptor word).
2	2	SMF104SEG	2	binary	Segment descriptor (see record length field).
4	4	SMF104FLG	1	binary	System indicator: **Bit** — **Meaning When Set** **0** — New record format **1** — Subtypes used **2** — Reserved **3-6** — Version indicators* **7** — System is running in PR/SM mode
5	5	SMF104RTY	1	binary	Record type 104 (X'68').
6	6	SMF104TME	4	binary	Time since midnight, in hundredths of a second, that the record was moved into the SMF buffer.
10	A	SMF104DTE	4	packed	Date when the record was moved into the SMF buffer, in the form 0*cyydddF*.
14	E	SMF104SID	4	EBCDIC	System identification (from the SMFPRMxx SID parameter).
18	12	SMF104SSI	4	EBCDIC	Subsystem identification (GPM).
22	16	SMF104STY	2	binary	Record subtype.
24	18	SMF104TRN	2	binary	Number of triplets in this record. A triplet is a set of three SMF fields (offset/length/number values) that defines a section of the record. The offset is the offset from the RDW.
26	1A		2		Reserved.
28	1C	SMF104PRS	4	binary	Offset to RMF XP product section from the RDW.
32	20	SMF104PRL	2	binary	Length of RMF XP product section.
34	22	SMF104PRN	2	binary	Number of RMF XP product sections.
Header extension for all subtypes					
36	24	SMF104ICS	4	binary	Offset to image control section from the RDW.
40	28	SMF104ICL	2	binary	Length of image control section.
42	2A	SMF104ICN	2	binary	Number of image control sections.
44	2C	SMF104MES	4	binary	Offset to metric section from the RDW.
48	30	SMF104MEL	2	binary	Length of metric section.
50	32	SMF104MEN	2	binary	Total number of metric sections.

- GPM4CIM writes one SMF record subtype per interval. That means that one subtype contains the performance data for all systems where GPM4CIM is connected to.
- Control sections will be provided to allow exploiters of the SMF records to assign the performance data to the appropriate systems.
- The stucture of the SMF record type 104 follows the general guidelines for SMF records including the common header.
- All sections can be addressed by means of so called triplets, where a triplet describes individual sections with offset, length and number.
- This allows to extract data from the new SMF record with standard SMF postprocessing tools.

RMF XP & SMF Records…

RMF Product Section with common Layout

Offsets		Name	Length	Format	Description
0	0	SMF104MFV	2	packed	RMF version number.
2	2	SMF104PRD	8	EBCDIC	Product name (***RMF XP***).
10	A	SMF104IST	4	packed	Time that the RMF XP measurement interval started, in the form 0*hhmmssF*, where *hh* is the hours, *mm* is the minutes, *ss* is the seconds, and *F* is the sign.
14	E	SMF104DAT	4	packed	Date when the RMF measurement interval started, in the form 0*cyydddF*.
18	12	SMF104INT	4	packed	Duration of RMF measurement interval, in the form *mmsstttF*, where *mm* is the minutes, *ss* is the seconds, *ttt* is the milliseconds, and *F* is the sign. The end of the measurement interval is the sum of the recorded start time and this field.
22	16	SMF104LGO	8	binary	Offset GMT to local time (STCK format).
30	1E		2		Reserved.
32	20	SMF104XPL	2	binary	RMF XP functionality level.
34	12	SMF104CPX	24	EBCDIC	System complex name, specified with the COMPLEX parameter in the *cfg4A/X/Z* configuration file.
58	3A	SMF104OSL	8	EBCDIC	Operating system label served by RMF XP (AIX or LINUX).
66	42	SMF104PLT	2	binary	Platform type served by RMF XP: **0** System p **1** System x **2** System z
68	44	SMF104MVS	8	EBCDIC	z/OS software level for the current system (consists of an acronym and the version, release, and modification level - ZV*vvrrmm*).
76	4C	SMF104XNM	8	EBCDIC	Sysplex name of the current sysplex as defined in parmlib member COUPLExx.
84	54	SMF104SNM	8	EBCDIC	System name for the current system as defined in parmlib member IEASYSxx SYSNAME parameter.

- The SMF 104 RMF Product Section is slightly different from the SMF 70-79 records since a subset of information is not available for the new platforms (e.g. cycle length, sample number).
- On the other hand, the SMF 104 RMF Product Section contains information which is not available in a z/OS environment (e.g. SMF104CPX, SMF104OSL, SMF104PLT).

RMF XP & SMF Records…

Image Control Section

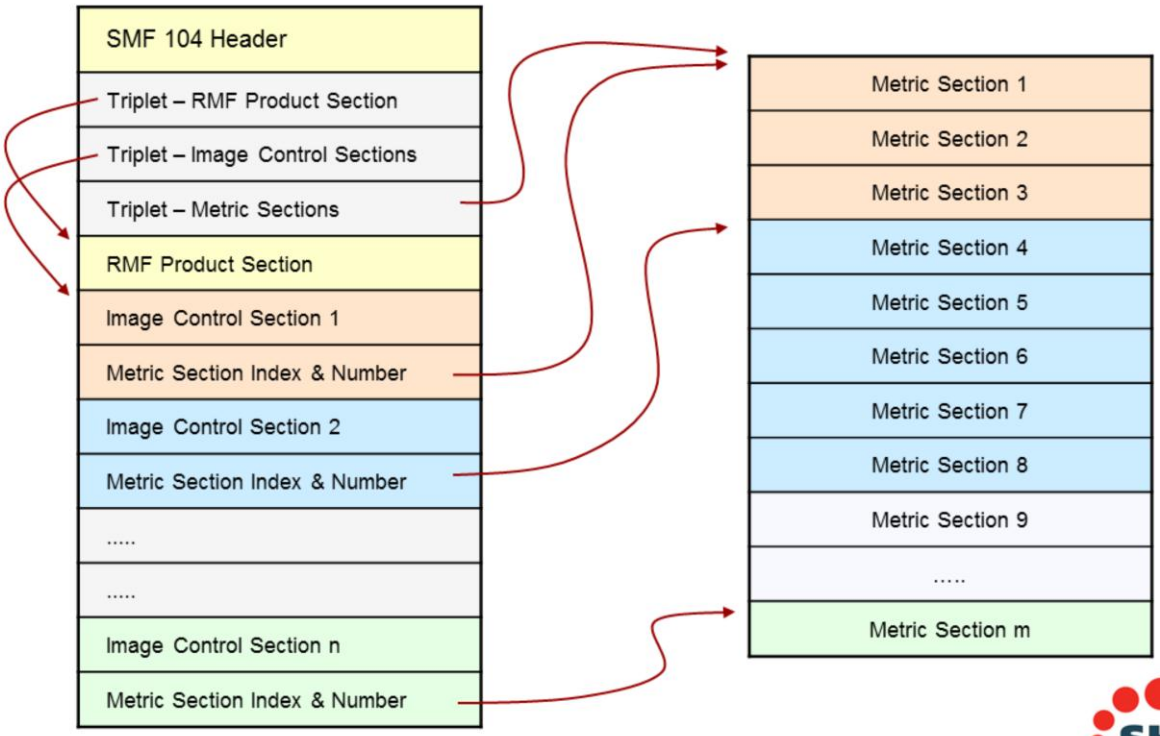
Offsets		Name	Length	Format	Description
0	0	SMF104MIM	64	EBCDIC	Name of this monitored image, extracted from the CIM metrics collection.
64	40	SMF104TIM	14	EBCDIC	Timestamp in the format *yyyymmddhhmmss*, extracted from the CIM metrics collection.
78	4E	SMF104DUR	14	EBCDIC	Interval duration in the format *yyyymmddhhmmss*, extracted from the CIM metrics collection.
92	5C	SMF104CIM	64	EBCDIC	Name of the image where the CIM server is running, specified with the IMAGE parameter in the *cfg4AIXIZ* configuration file.
156	9C	SMF104OST	4	EBCDIC	Operating system type where the CIM server is running, extracted from the OSType attribute of the CIM_Operating_System instance: **9** AIX **36** Linux
160	A0	SMF104OSV	64	EBCDIC	Operating system version where the CIM server is running, extracted from the version attribute of the CIM_Operating_System instance.
224	E0	SMF104CTZ	4	EBCDIC	Current time zone, extracted from the CurrentTimeZone attribute of the CIM_Operating_System instance. This value represents the GMT offset in minutes.
228	E4	SMF104MIND	2	binary	Index of first metric section associated with this monitored image.
230	E6	SMF104MNUM	2	binary	Number of metric sections associated with this monitored image.

Metric Section

Offsets		Name	Length	Format	Description
0	0	R10406MNAME	64	EBCDIC	Name of measured element, extracted from the MeasuredElementName attribute of the CIM_BaseMetricValue instance.
64	40	R10406AS	8	floating	Available space for this filesystem in megabytes. (AvailableSpace)
72	48	R10406TS	8	floating	Total space for this filesystem in megabytes. (TotalSpace)
80	50	R10406US	8	floating	Used space for this filesystem in megabytes. (UsedSpace)

- The SMF Type 104 Record contains performance data that has been collected from one or more systems resp. images.
- There is one image control section per image. Beyond general information (e.g. time stamps) the image control section describes the corresponding metric sections with the measured elements and their metric values.

RMF XP & SMF Records…



- All image control sections are located consecutively right behind the RMF product section.
- In the same way, all metric sections are following right behind the last image control section.
- The metric section index and number within the image control section allows to identify the metric sections belonging to a certain image.
- Since one SMF record is limited to the length of 32756 bytes, an individual record can be splitted and not all metric sections for a certain image might be located in the same record.
- However, in any case, one single record is self-containing: that means, a meaningful report can be produced just by using one single record. There is no need to combine multiple records from the same subtype for reporting purposes.

RMF XP & SMF Records…

RMF ERBSCAN Utility maps all Sections of the SMF Type 104 record

```
-> AIX Image Control Section (2)
================================

   #1:   +0000:  97F69994 86F10000 00000000 00000000  *p6rmf1          *
         +0010:  00000000 00000000 00000000 00000000  *                *
         +0020:  00000000 00000000 00000000 00000000  *                *
         +0030:  00000000 00000000 00000000 00000000  *                *
         +0040:  F2F0F1F1 F0F9F0F8 F1F6F2F7 F0F3F0F0  *2011090816270300*
         +0050:  F0F0F0F0 F0F0F0F0 F0F1F0F1 97F69994  *000000000101p6rm*
         +0060:  86F14B82 96858293 89958785 954B8485  *f1.boeblingen.de*
         +0070:  4B898294 4B839694 00000000 00000000  *.ibm.com        *
         +0080:  00000000 00000000 00000000 00000000  *                *
         +0090:  00000000 00000000 00000000 F9000000  *            9   *
         +00A0:  F64BF14B F04BF000 00000000 00000000  *6.1.0.0         *
         +00B0:  00000000 00000000 00000000 00000000  *                *
         +00C0:  00000000 00000000 00000000 00000000  *                *
         +00D0:  00000000 00000000 00000000 00000000  *                *
         +00E0:  F6F00000 00000002                    *60              *

      -> Image name        : p6rmf1
      -> Operating system  : 6.1.0.0
      -> First metric sec. : #1
      -> Number of metrics : 2

-> AIX_ProcessorMetrics (4)
===========================

   #1:   +0000:  8397A4F0 00000000 00000000 00000000  *cpu0            *
         +0010:  00000000 00000000 00000000 00000000  *                *
         +0020:  00000000 00000000 00000000 00000000  *                *
         +0030:  00000000 00000000 00000000 00000000  *                *
         +0040:  414C5441 355475A3 425F383E 425AEE63  * <è  èÎtâ^  â!ÓÄ*
         +0050:  412E1298 88F861A6 411E41B3 28B6D86F  *   qh8/w   · ¶Q?*
         +0060:  3F28064E A3A70EA2                    *   +tx s        *

      -> Measured element : cpu0
      -> List of values   :
                            '414C5441355475A3'X   4.77057000
                            '425F383E425AEE63'X   95.2196999
                            '412E129888F861A6'X   2.87954000
                            '411E41B328B6D86F'X   1.89104000
                            '3F28064EA3A70EA2'X   0.00977164000
```


- The RMF utility ERBSCAN has been extended to support the new SMF type 104
- Beyond the header sections, the content of the image sections as well as the metric sections is displayed in hex and in character format.
- The individual metric values are stored in double floating point format. ERBSCAN translates all the values to the decimal format and displays it right behind the hexdump section.

RMF XP & SMF Records…

```
//GPM4CIM  PROC OS=A
//STEP1    EXEC PGM=BPXBATCH,TIME=NOLIMIT,REGION=0M,
//       PARM='PGM /usr/lpp/gpm/bin/gpm4cim cfg=/etc/gp /gpm4&OS..cfg'
//STDENV   DD   PATH='/etc/gpm/gpm4cim.env'
//STDOUT   DD   PATH='/var/gpm/logs/gpm4cim&OS..out',
//         PATHOPTS=(OWRONLY,OCREAT,OTRUNC),
//         PATHMODE=(SIRUSR,SIWUSR,SIRGRP)
//STDERR   DD   PATH='/var/gpm/logs/gpm4cim&OS..trc',
//         PATHOPTS=(OWRONLY,OCREAT,OTRUNC),
//         PATHMODE=(SIRUSR,SIWUSR,SIRGRP)
//SYSPRINT DD   SYSOUT=*
//SYSOUT   DD   SYSOUT=*
//         PEND
```

```
MAXSESSIONS_HTTP(20)                          /* MaxNo of concurrent
HTTP_PORT(8805)                               /* Port number for HT
HTTP_ALLOW(*)                                 /* Mask for hosts tha
HTTP_NOAUTH(*)                                /* No server can acces
INTERVAL(300)                                 /* Length of the monitor
AIX_COMPLEX(WEBPLEX)                          /* Name of system complex
AIX_IMAGE(p6rmf1.boeblingen.de.ibm.com:5988)  /* Hostname of member  */
AIX_IMAGE(p6rmf2.boeblingen.de.ibm.com:5988)  /*                     */
RECORD                                        /* Write SMF Records   */
```

New global Option:
RECORD / NORECORD
Default: NORECORD

Change RECORD Option dynamically: F GPM4CIM, RECORD/NORECORD

✓ GPM253I SMF RECORDING IS NOW ON/OFF

- A new option RECORD/NORECORD within the RMF XP configuration file (/etc/gpm/gpm4[A|X|Z|W].cfg) is introduced to specify that the GPM4CIM server writes SMF type 104 records
- The scope of this option is global, this means that the option is applied to all supported subtypes.
- Of course, SMF recording can also be enabled or disabled with a MODIFY command:

  F GPM4CIM,[NO]RECORD.

- Once the RECORD parameter is set to on, SMF 104 records are written frequently and the completion of an interval is indicated by new message:

  GPM280I RMF XP INTERVAL PROCESSING FOR name COMPLETE

RMF XP & SMF Records…

```
ACTIVE                          /* ACTIVE SMF RECORDING*/
DSNAME(SYS1.&SYSNAME..MAN1,
       SYS1.&SYSNAME..MAN2,
       SYS1.&SYSNAME..MAN3)
MAXDORM(3000)                   /* WRITE AN IDLE BUFFER
MEMLIMIT(20000M)                /* MEMLIMIT ABOVE THE BA
STATUS(010000)                  /* WRITE SMF STATS AFTER
JWT(0900)                       /* 522 AFTER 15 HOURS
SID(&SMFID)                     /* SYSTEM ID IS SYSBLD
LISTDSN                         /* LIST DATA SET STATUS AT IPL  */
INTVAL(15)
SYNCVAL(00)
SYS(TYPE(30,42,70:79,103,104(1:12,20:31,40:53,60:64),108),
    EXITS(IEFU83,IEFU84,IEFU85,IEFACTRT,IEFUJV,IEFUSI,
    IEFUJP,IEFUSO,IEFUJI,IEFUTL,IEFU29,IEFUAV),
    INTERVAL(SMF,SYNC),NODETAIL)
```

Control SMF Recording on Subtype Level via SMFPRMxx Parmlib Member

SMFWTM → SMF Type 104

```
//RMF       PROC
//IEFPROC   EXEC PGM=ERBMFMFC,REGION=32M,TIME=1440,
//          PARM='SMFBUF(RECTYPE(30,70:79,104(1:12,20:31,40:53,60:64)))'
```

SMFWTM → SMF Buffer of RMF Sysplex Data Server

Control SMF Buffering on Subtype Level via RMF and SMFBUF Parameter

- Use the SMF built-in facilities in order to control the SMF 104 recording on subtype level. You can dynamically switch on or off SMF recording for a specific subtype by changing the active SMFPRMxx parmlib member. For example:

  SYS(TYPE(30,42,70:79,103,104(1:12,20:31,40:53,60:64),108),…

- The capability to receive and hold real-time copies of SMF records in the in-storage SMF buffer of the RMF Sysplex Data Server is a standard feature of RMF. To use this feature, modify the PARM statement of the RMF procedure to include SMF type 104.

RMF Enhancements at a Glance

- IBM zEnterprise EC12 Support
  - Statistics for Flash Memory and Pageable Large Pages
  - Support of Coupling Facility Flash Memory
  - z Enterprise Data Compression Express Reporting (zEDC)
  - Shared Memory Communication Reporting (SMC-R)
  - Support of I/O Interrupt Delay Time Facility
  - Extended Infiniband Link Reporting
  - Support of Crypto Express4 Card

- IBM z13 Support
  - RMF support for SMT, Crypto Express5S …
  - PCIe reporting enhancements

- Exploitation of System z Integrated Information Processors
  - zIIP Usage Option for Monitor III Gatherer

- Postprocessor XML formatted Reports
  - Transition to XML Format almost complete
  - Advanced Sorting and Filtering Capabilities

- RMF XP Enhancements
  - SMF Type 104 Recording for all Platforms

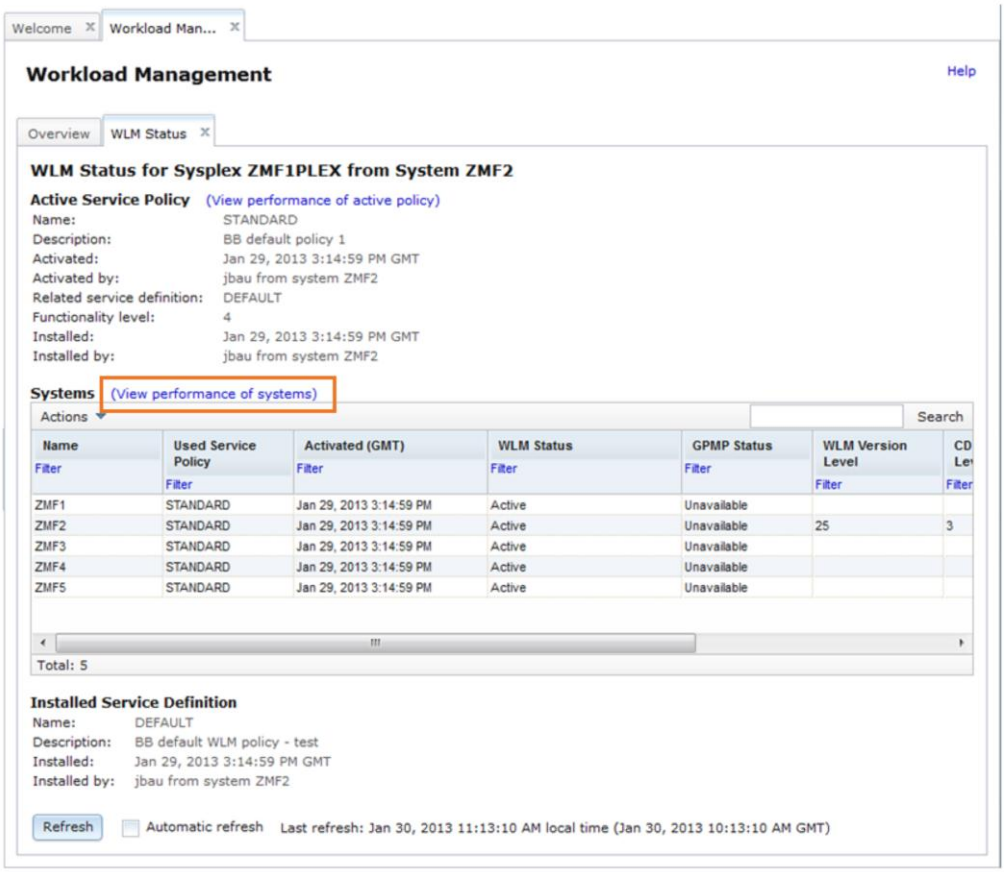
- z/OSMF Resource Monitoring
  - Context sensitive Application Linkage to WLM
  - Historical Reporting & Spreadsheet Export

z/OSMF Application Linking (Resource Monitoring & WLM)

- The definitions of Workload Management determine the performance behavior of the systems.
- Resource Monitoring visualizes the performance behavior.
- Link z/OSMF WLM and RM to each other:
  - When you work with WLM service definitions: Jump to Resource Monitoring to visualize the resulting performance.
  - When you detect abnormal metric values in Resource Monitoring: Jump to Workload Management to look at the service definition.
- Performance metrics can be viewed more easily in context with the active service definition/policy and vice versa.

- The new z/OSMF Application Linking allows to link the following tasks of the Performance category to each other in context:
  - The z/OSMF Workload Managemetn task, which allows you to define, manage, install and activate your WLM service definitions and policies.
  - The z/OSMF System Status, which provides a quick performance status overview of all the sysplexes in your data center.
  - The z/OSMF Resource Monitoring task, which allows you to define monitoring dashboards that show performance results in detailed bar charts.
- z/OSMF Application Linking support is available for z/OS 1.13 with the following new function APARs:
  - PM74502 (z/OSMF Base)
  - PM74508 (Resource Monitoring)
  - PM74517 (WLM)
  - PM74925 (WLM)
- The following foils show examples of links between the z/OSMF WLM and Resource Monitoring plug-ins.

From WLM Status to RM System Status



- By linking the Workload Management task and the System Status task of the z/OSMF Resource Monitoring plug-in, you can opne the System Status task from the Workload Management task and view the status of the sysplex.
- From System Status task you can open the Workload Management task and view the active service definition, active service policy or WLM status.

RM System Status

Welcome | Workload Man... | System Status

Help

System Status

Use this page to quickly assess the performance of the workloads running on the sysplexes in your installation. You can also use this page to define the target systems for the sysplexes and AIX or Linux system complexes that you want to monitor in the Resource Monitoring task.

Resources

Actions

Resource	System Type	Connectivity	Performance Index Status	Related Service Definition	Active WLM Policy
Filter	Filter	Filter	Filter	Filter	Filter
LOCALPLEX	z/OS	Connected	PI > 1 for unimportant periods	DEFAULT	STANDARD
SCLMPLEX	z/OS	Connected	PI <= 1 for all periods	Default	STANDARD
SYSPLEX	z/OS	Connected	PI <= 1 for all periods	SYSTES2	STANDARD
IRDPLEX	z/OS	Error			

Total: 4, Selected: 1

Refresh Last refresh: Jan 30, 2013 11:18:00 AM local time (Jan 30, 2013 10:18:00 AM GMT)

Automatic refresh

From RM System Status

Welcome | Workload Man... | System Status

Help

System Status

Use this page to quickly assess the performance of the workloads running on the sysplexes in your installation. You can also use this page to define the target systems for the sysplexes and AIX or Linux system complexes that you want to monitor in the Resource Monitoring task.

Resources

Actions

Resource	System Type	Connectivity	Performance Index Status	Related Service Definition	Active WLM Policy
Filter	Filter	Filter	Filter	Filter	Filter
LOCALPLEX	z/OS	Connected	PI > 1 for unimportant periods	DEFAULT	STANDARD
SCLMPLEX		Connected	PI <= 1 for all periods	Default	STANDARD
SYSPLEX			PI <= 1 for all periods	SYSTES2	STANDARD
IRDPLEX	z/OS				

Modify Entry
Remove Entry
View
- Performance Index Details
- Active WLM Service Definition
- Active WLM Policy
- WLM Status

Total: 4, Selected: 1

Refresh Last refresh: Jan 30, 2013 11:20:04 AM local time (Jan 30, 2013 10:20:04 AM GMT)

Automatic refresh

From RM System Status

Welcome | Workload Man... | System Status

Help

System Status

Use this page to quickly assess the performance of the workloads running on the sysplexes in your installation. You can also use this page to define the target systems for the sysplexes and AIX or Linux system complexes that you want to monitor in the Resource Monitoring task.

Resources

Actions

Resource	System Type	Connectivity	Performance Index Status	Related Service Definition	Active WLM Policy
Filter	Filter	Filter	Filter	Filter	Filter
LOCALPLEX	z/OS	Connected	PI > 1 for unimportant periods	DEFAULT	STANDARD
SCLMPLEX		Connected	PI <= 1 for all periods	Default	STANDARD
SYSPLEX			PI <= 1 for all periods	SYSTES2	STANDARD
IRDPLEX	z/OS				

Modify Entry
Remove Entry
View
- Performance Index Details
- Active WLM Service Definition
- Active WLM Policy
- WLM Status

Resource Monitoring

Dashboards | Performance Index - LOCALPLEX

Performance Index - LOCALPLEX (Running)

Start Pause Save Actions

Important Service Class Periods		Service Class Periods	
PRDTSO.1	0	STCCMD.1	0.67
		PRDTSO.1	0

Total: 4, Selected: 1

Refresh Last refresh: Jan 30, 2013 11:20:04 AM local time (Jan 30, 2013 10:20:04 AM GMT)

Automatic refresh

From RM System Status

Welcome | Workload Man... | System Status

Help

System Status

Use this page to quickly assess the performance of the workloads running on the sysplexes in your installation. You can also use this page to define the target systems for the sysplexes and AIX or Linux system complexes that you want to monitor in the Resource Monitoring task.

Resources

Actions

Resource	System Type	Connectivity	Performance Index Status	Related Service Definition	Active WLM Policy
Filter	Filter	Filter	Filter	Filter	Filter
LOCALPLEX	z/OS	Connected	PI > 1 for unimportant periods	DEFAULT	STANDARD
SCLMPLEX		Connected	PI <= 1 for all periods	Default	STANDARD
SYSPLEX			PI <= 1 for all periods	SYSTES2	STANDARD
IRDPLEX	z/OS				

Modify Entry
Remove Entry
View
- Performance Index Details
- Active WLM Service Definition
- Active WLM Policy
- WLM Status

Resource Monitoring

Dashboards | Performance Index - LOCALPLEX

Workload Management

Overview | WLM Status | View DEFAULT

This service definition is installed and policy STANDARD is active

Service Definition Details

Service definition name:
DEFAULT

Description:
BB default WLM policy - test

Functionality level:
004

Total: 4, Selected: 1

Refresh Last refresh: Jan 30, 2013 11:20:04 AM local time (Jan 30, 2013 10:20:04 AM GMT)

Automatic refresh

From RM System Status

Welcome | Workload Man... | System Status

Help

System Status

Use this page to quickly assess the performance of the workloads running on the sysplexes in your installation. You can also use this page to define the target systems for the sysplexes and AIX or Linux system complexes that you want to monitor in the Resource Monitoring task.

Resources

Actions

Resource	System Type	Connectivity	Performance Index Status	Related Service Definition	Active WLM Policy
Filter	Filter	Filter	Filter	Filter	Filter
LOCALPLEX	z/OS	Connected	PI > 1 for unimportant periods	DEFAULT	STANDARD
SCLMPLEX		Connected	PI <= 1 for all periods	Default	STANDARD
SYSPLEX			PI <= 1 for all periods	SYSTES2	STANDARD
IRDPLEX	z/OS				

Modify Entry
Remove Entry
View
- Performance Index Details
- Active WLM Service Definition
- Active WLM Policy
- WLM Status

Total: 4, Selected: 1

Refresh Last refresh: Jan 30, 2013 11:20:04 AM local time (Jan 30, 2013 10:20

Automatic refresh

Resource Monitoring

Workload Management

Workload Management

Overview | WLM Status | View DEFAULT

Service Policies ▸ Properties This service definition is installed and policy STANDARD is active

Properties for Active Service Policy

Service policy name: STANDARD

Description: BB default policy 1

Service Class Overrides | Resource Group Overrides

Actions Table view: Tree

Service Class	Period	Importance	Duration	Goal Type	Resp Goal
Filter	Filter	Filter	Filter	Filter	Filter

From RM System Status

Welcome | Workload Man... | System Status

Help

System Status

Use this page to quickly assess the performance of the workloads running on the sysplexes in your installation. You can also use this page to define the target systems for the sysplexes and AIX or Linux system complexes that you want to monitor in the Resource Monitoring task.

Resources

Actions

Resource	System Type	Connectivity	Performance Index Status	Related Service Definition	Active WLM Policy
Filter	Filter	Filter	Filter	Filter	Filter
LOCALPLEX	z/OS	Connected	PI > 1 for unimportant periods	DEFAULT	STANDARD
SCLMPLEX		Connected	PI <= 1 for all periods	Default	STANDARD
SYSPLEX			PI <= 1 for all periods	SYSTES2	STANDARD
IRDPLEX	z/OS				

Modify Entry
Remove Entry
View
- Performance Index Details
- Active WLM Service Definition
- Active WLM Policy
- WLM Status

Total: 4, Selected: 1

Refresh Last refresh: Jan 30, 2013 11:20:04 AM local time (Jan 30, 2013 10:20

Automatic refresh

Resource Monitoring

Workload Management

Workload Management

Overview | WLM Status | View DEFAULT

ANDARD is active

Workload Management

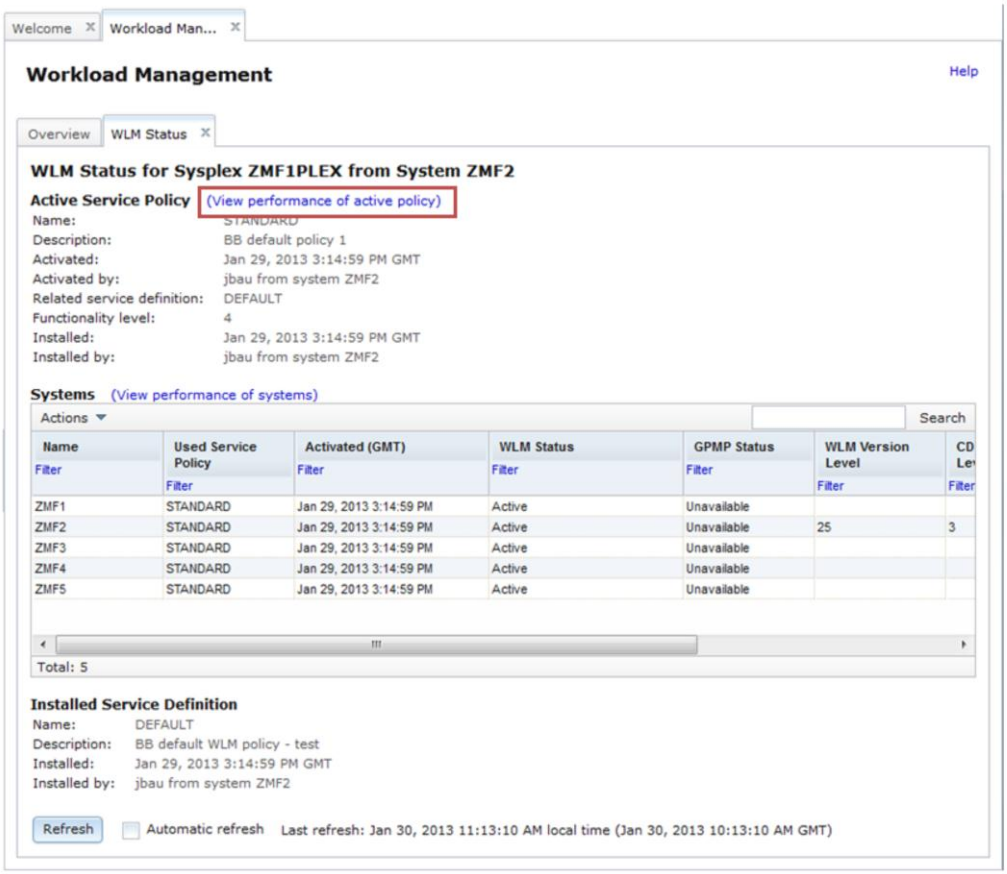
Overview | WLM Status

WLM Status for Sysplex ZMF1PLEX from System ZMF2

Active Service Policy (View performance of active policy)

Name:	STANDARD
Description:	BB default policy 1
Activated:	Jan 29, 2013 3:14:59 PM GMT
Activated by:	jbau from system ZMF2
Related service definition:	DEFAULT
Functionality level:	4
Installed:	Jan 29, 2013 3:14:59 PM GMT
Installed by:	jbau from system ZMF2

From WLM to RM Dashboard (Service Classes)



- By linking the Workload Management task and the Resource Monitoring task, you can open dashboards with performance metrics for service classes, workloads and report classes from within the Workload Management task.
- From a Resource Monitoring dashboard with performance metrics for service classes, workloads and report classes you can open the corresponding information in the WLM service definition.

RM Dashboard – WLM Service Class Performance

Welcome | Workload Man... | System Status | Resource Mon...

Help

Resource Monitoring

Dashboards | WLM Service Class - LOCALPLEX

WLM Service Class - LOCALPLEX (Running)

Start Pause Save Actions

Performance Index

STCCMD. 0.56
HOTTSO. 0.5

,LOCALPLEX,SYSPLEX *performance index by WLM service class period*

01/30/2013 11:24:00 - 01/30/2013 11:25:00 (3/3)

Execution Velocity

STCCMD. 71 40
IRLM. 0 75
OMVS. 0 20
OMVS. 0 30

,LOCALPLEX,SYSPLEX *execution velocity by WLM service class period*
,LOCALPLEX,SYSPLEX *execution velocity goal by WLM service class period*

01/30/2013 11:24:00 - 01/30/2013 11:25:00 (3/3)

Response Time

HOTTSO. 0.015 N/A
TSOODD. 0 0.3
TSOODD. 0 1.5

Percentile Response Time

HOTTSO. 100 95
TSOMED. N/A 90
TSOREG. N/A 90

From RM Dashboard to WLM Service Classes

Welcome | Workload Man... | System Status | Resource Mon...

Help

Resource Monitoring

Dashboards | WLM Service Class - LOCALPLEX

WLM Service Class - LOCALPLEX (Running)

Start Pause Save Actions

Performance Index

STCCMD.1 0.56

HOTTSO.1 0.5

View Resource Attributes

View WLM Service Class HOTTSO

,LOCALPLEX,SYSPLEX *performance index by WLM service class period*

01/30/2013 11:24:00 - 01/30/2013 11:25:00 (3/3)

Execution Velocity

STCCMD.1 71 40

IRLM.1 0 75

OMVS.2 0 20

OMVS.1 0 30

,LOCALPLEX,SYSPLEX *execution velocity by WLM service class period*

,LOCALPLEX,SYSPLEX *execution velocity goal by WLM service class period*

01/30/2013 11:24:00 - 01/30/2013 11:25:00 (3/3)

Response Time

HOTTSO.1 0.015 N/A

TSOODD.1 0 0.3

TSOODD.2 0 1.5

Percentile Response Time

HOTTSO.1 100 95

TSOMED.2 N/A 90

TSOREG.2 N/A 90

WLM Service Classes

Welcome | Workload Man... | System Status | Resource Mon...

Workload Management Help

Overview | WLM Status | View DEFAULT

This service definition is installed and policy STANDARD is active Notes Switch To

Service Classes

Actions Table view: Tree Search

Name	Period	Importa	Durati	Goal Type	Response Time Goal (hh:mm:ss.ttt)	Percent Goal	Velocity Goal	CPU Critical	Resource Group	Workloa
Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
BATCH1								No		BATCI
BATCH2								No		BATCI
BATCHHI								No		BATCI
BATCHLOW								No		BATCI
BATCHMED								No		BATCI
BATCHRSP								No		BATCI
DISCRET								No		BATCI
HOTTSO								No	REGTSO	TSO
HOTTSO	1	1		Percentile Response Time	00:00:00.500	95			REGTSO	TSO
IRLM								No		BATCI
OE								No	REGTSO	OMVS
OMVS								No		BATCI
OMVSKERN								No		BATCI
PRDTSO								No	REGTSO	TSO
STCCMD								No	REGSTC	STC
STCLO								No	BATCHVEL	STC
STCSYS								No	HIGHPRTY	STC
STORPROC								No		BATCI
TSOEVEN								No		BATCI

Total: 62, Selected: 1

From WLM to RM Dashboard (Workloads)

Welcome | Workload Man... | System Status | Resource Mon...

Workload Management Help

Overview | WLM Status | Modify DEFAULT

This service definition is installed and policy STANDARD is active Notes Switch To

Workloads

Actions Search

Name	Messages	Last Modified (GMT)	Modified By
Filter	Filter	Filter	Filter
APPC		Mar 30, 2011 2:31:17 PM	debug1
BATCH		Oct 16, 1998 11:58:16 AM	tage
CICS		Oct 16, 1998 11:58:30 AM	tage
OMVS		Oct 16, 1998 12:01:03 PM	tage
STC		Oct 16, 1998 12:01:45 PM	tage
TSO		Oct 16, 1998 12:01:52 PM	tage

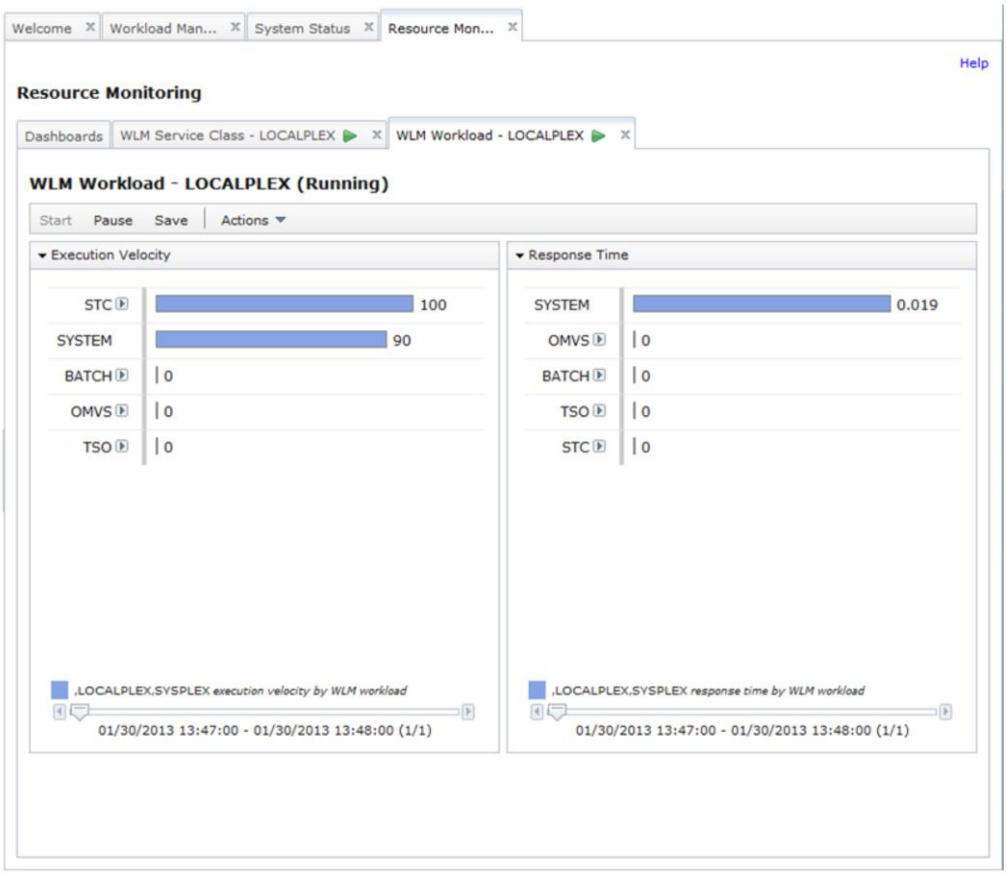
Cut to Clipboard
Copy to Clipboard
Delete...
View Cross References
View Messages
View Performance of Selected
New...
Paste
View Performance of All
Select All
Deselect All
Configure Columns...
Modify Filters...
Hide Filter Row
Clear Filters
Modify Sort...
Clear Sorts
Clear Search

Total: 6, Selected: 1

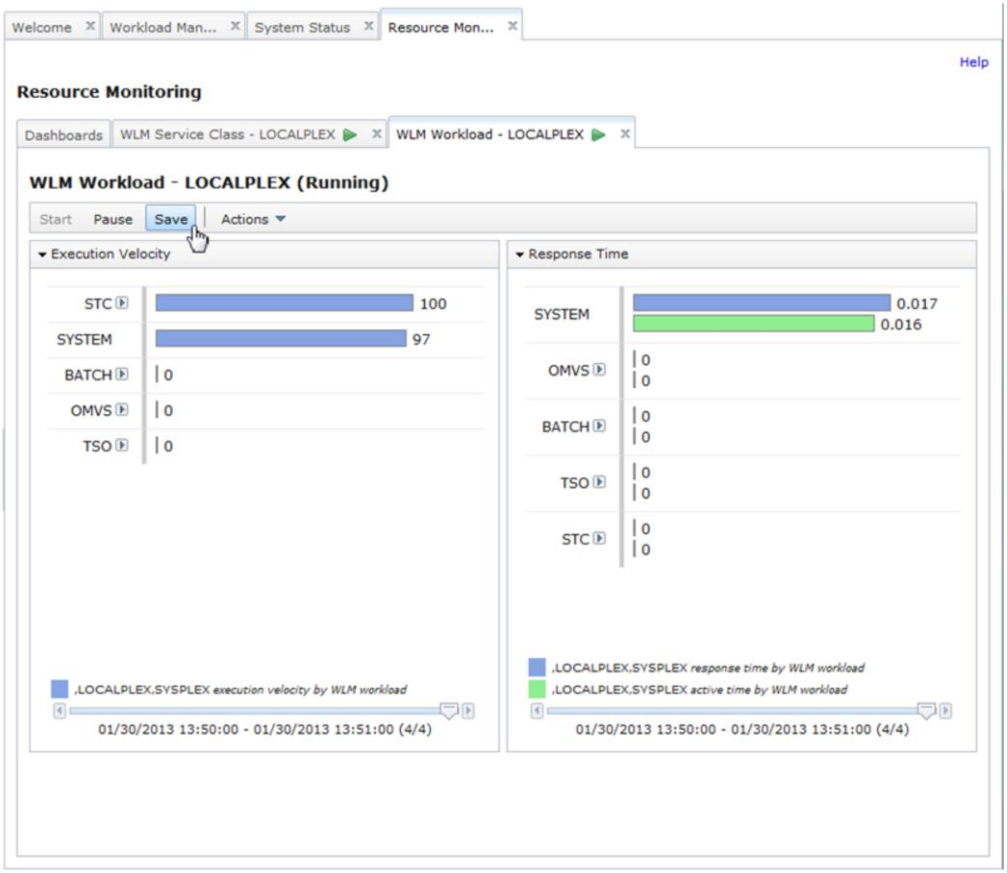
Reapply Filter and Sort

OK Apply Reset Cancel

RM Dashboard – WLM Workload Performance



Customization and Persistence



- The user can customize dashboards opened by Application Linking and save them to the Dashboards list.
- Then the Dashboard can be opened directly in Resource Monitoring using the Dashboards list.
- Subsequent application linking events will use the saved dashboard.

Conditions for Application Linking between WLM & RM

- In the Workload Management task, the *View Performance…* actions and links are only available if the service definition in the *View/Modify* tab is currently activated in the Sysplex.
- In the System Status task, the WLM related *View* actions (and corresponding links) are only available if the selected resource is the z/OS sysplex where z/OSMF is running in (local sysplex).
- In a monitoring dashboard, the context menu icon is only visible if the performance data is retrieved from the local sysplex and the chart is related to WLM definitions, i.e.,
  - The resources in the chart are WLM service classes, service class periods, report classes, or workloads.
  - The metric is filtered by a workscope of a WLM service class, service class period, report class, or workload.
    (Example: *% using by MVS image [BATCH,S]* , where *[BATCH,S]* means: filtered by workscope of service class *BATCH*)

Resource Monitoring – Historical Reporting

Welcome | Resource Mon...

Resource Monitoring

Dashboards | Storage Soaker

Storage Soaker (Running)

Start Pause Stop Save Actions

Active & Fixed Frames (WLM Vie... | ...ve Frames

SYSTEM.1 2907 | FTPS1 [0041] 564

SYSSTC.1 4436 | CBIND1 [005D] 499

STCDEF.1 3530 203000

GPMSERVE.1 261 428

OE.2 0 564

OMVSKERN.1 0 390

SYSF,*,STORAGE # frames fixed by WLM service class period

SYSF,*,STORAGE # frames active by WLM service class period

05/02/2014 10:35:00 - 05/02/2014 10:36:00 (1/1)

Start Dashboard
Pause Dashboard
Retrieve Historical Data...
Stop Retrieving Historical Da...
Save Dashboard
Rename Dashboard...
Modify Settings...
Add Metric...
Modify Metric...
Export...
Delete Metric or Group
Arrange Metrics or Gro...
Rename Metric Group...

Retrieve Historical Data

To retrieve historical data from the DDS, specify the metric groups, timeframe, and range. This action may generate significant amount of overhead on the DDS host, especially for long time ranges.

Metric groups:

- Active & Fixed Frames (WLM View)
- Active Frames
- Fixed Frames

Timeframe:

Condition:	Amount:	Unit:
Past	2	Hours

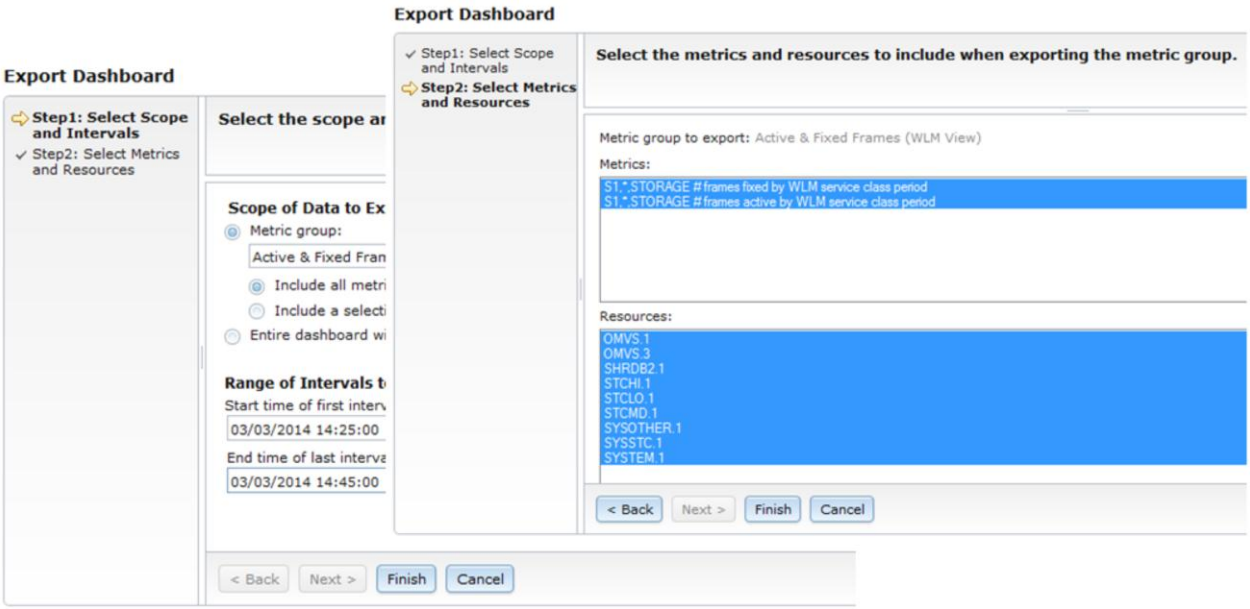
Data sample range:

- Use the default range
- Specify the range in seconds: 600

OK | Restore Defaults | Cancel | Help

Retrieve performance data from the past accordingly to the content of the RMFGAT in-storage buffer or the RMFGAT VSAM Data Sets

Resource Monitoring – Spreadsheet Export



Information and Tools

- RMF website: www.ibm.com/systems/z/os/zos/features/rmf/
  - Product information, newsletters, presentations, ...
  - Downloads
    - Spreadsheet Reporter
    - RMF PM Java Edition
    - Postprocessor XML Toolkit

- RMF email address: rmf@de.ibm.com

Users Guide:
New RMF
XP Chapter

- Documentation and news:
  - RMF Performance Management Guide, SC33-7992
  - RMF Report Analysis, SC33-7991
  - RMF User's Guide, SC33-7990
  - Latest version of PDF files can be downloaded from:
    www.ibm.com/systems/z/os/zos/bkserv/v2r1pdf/#rmf

Function Reference

Function	Availability
Flash Memory & Pageable Large Pages	APAR OA38660
Support of Coupling Facility Flash Memory	APAR OA40515
zEnterprise Data Compression Express	z/OS 2.1 RMF
Shared Memory Communication Reporting	z/OS 2.1 RMF
I/O Interrupt Delay Time	APAR OA39993
IFB Link Reporting	APAR OA37826
Crypto Express5S Statistics	APAR OA37016
Support for SMT	APAR OA44101
PCIE enhancements for RoCE and zEDC	APAR OA44524
Crypto Express5S Statistics	APAR OA43493
Support for LPARs with up to 4 TB real storage	APAR OA44503
Support for z13 IBM Integrated Coupling Adapter (ICA SR)	APAR OA44502
Postprocessor XML Formatted Reports	z/OS 1.11 RMF – z/OS 2.1 RMF
z/OSMF Resource Monitoring – Application Linking	z/OS 2.1
z/OSMF Resource Monitoring – Historical Reporting	APAR PI08825
z/OSMF Resource Monitoring – Spreadsheet Export	APAR PI08825
RMF XP – SMF 104 Recording Facility	z/OS 2.1 RMF

zEC12 Enhancements

z13 Enhancements