z/OS
Resource Measurement Facility

The Latest and Greatest:
z/OS V2R2

November, 6th 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.
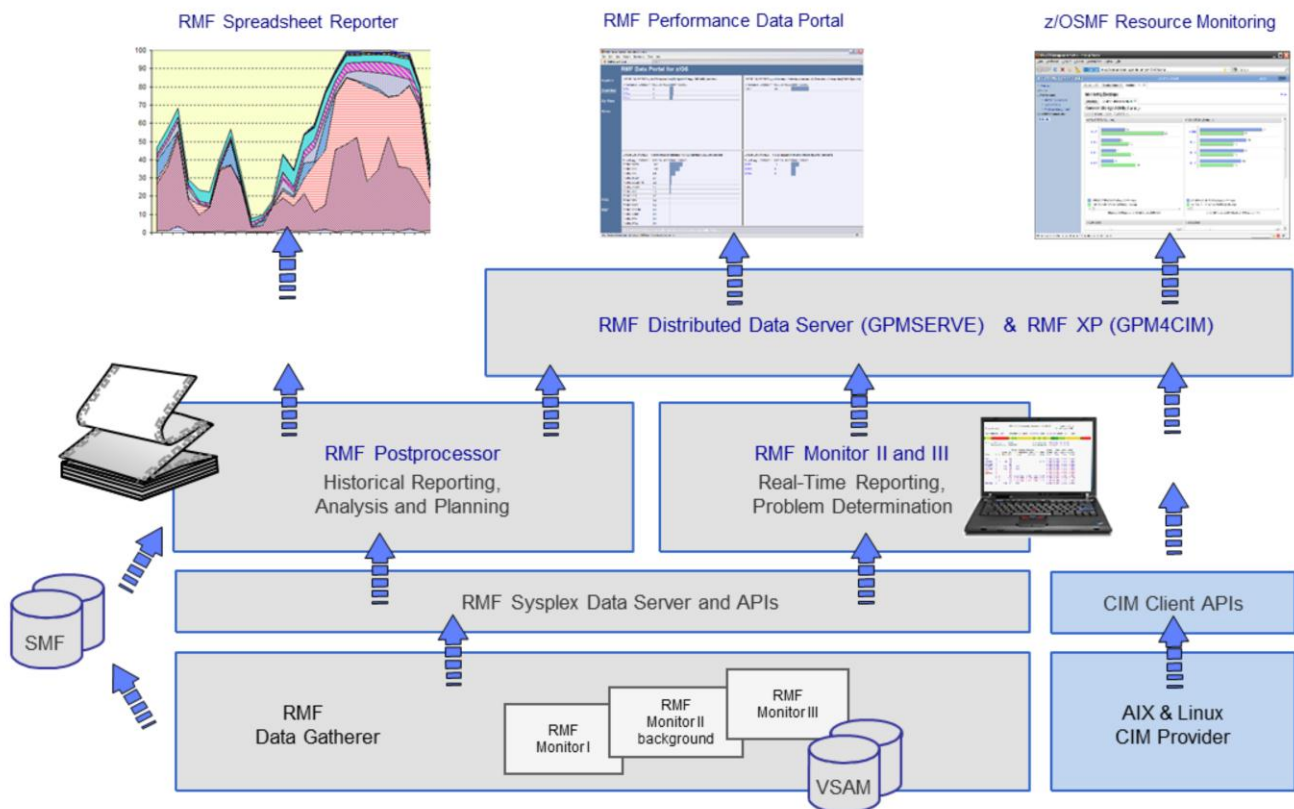This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available 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

RMF Spreadsheet Reporter | RMF Performance Data Portal | z/OSMF Resource Monitoring

RMF Distributed Data Server (GPMSERVE) & RMF XP (GPM4CIM)

RMF Postprocessor
Historical Reporting, Analysis and Planning

RMF Monitor II and III
Real-Time Reporting, Problem Determination

RMF Sysplex Data Server and APIs

CIM Client APIs

SMF

RMF Data Gatherer

RMF Monitor I

RMF Monitor II background

RMF Monitor III

VSAM

AIX & Linux CIM Provider

- 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 z13 Support**
  - ► RMF Statistics for Simultaneous Multithreading (SMT)
  - ► Extended ICSF Measurements for Crypto Express5S
  - ► Support for LPARs with up to 4TB Real Storage

- **SCM I/O Adapter Performance Reporting**
  - ► Statistics for EADM Subchannel Activities
  - ► Performance and Thruput on SCM Card Level

- **PCIE Activity Reporting**
  - ► z/OS V2.1 Postprocessor PCIE reporting
  - ► Additional Measurements for PCIe attached RoCE and zEDC Devices
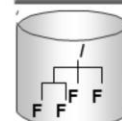 - ► z/OS V2.2 Monitor III PCIE Report

- **z/OS V2.2 Monitor III Job Resource Consumption Reporting**
  - ► CPU, I/O and Storage Consumption Data at a Glance
  - ► Detailed Statistics for Job related GQSCAN Activities

- **z/OS V2.2 zFS Reporting Enhancements**
  - ► Sysplex-wide Statistics for zFS Usage and Performance
  - ► Improved zFS Data Gathering Performance
  - ► Additional zFS Statistics for Shared File System environments

- **z/OSMF Resource Monitoring**
  - ► Historical Reporting & Spreadsheet Export

- **The zEvent Mobile Application**
  - ► Receive Push Messages based on critical System Events
  - ► Access to RMF Data Portal and z/OSMF Resource Monitoring

In accordance with the availability of new z/OS releases and new hardware functionality, the capabilities of RMF are enhanced Consecutively

With the availability of the IBM z13 servers, RMF provides first day support for a couple of notable hardware features.
- Comprehensive Statistics for Simultaneous Multithreading (SMT)
- Extended ICSF Measurements for Crypto Express5S
- Support for LPARs with up to 4TB Real Storage

Storage Class Memory – aka Flash Memory – is a new tier within the memory hierarchy of the zSeries family.
RMF provides detailed usage statistics for Storage Class Memory related operations by means of a new Monitor III SCM Activity report.
Together with the z EC12 server family two new types of PCIe cards have been introduced:
- The new z Enterprise Data Compression (zEDC) Express offering provides low-cost data compression for z/OS system services and applications
- 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 (similar than Hipersockets for LPAR to LPAR Communication)
- The existing Postprocessor PCIe Report has been extended with additional measurements
- With z/OS V2R2 RMF a complete new Monitor III PCIe Report allows to keep track of PCIe related operations in online mode

The Monitor III Job Usage Report complements the Monitor III reporting suite with detailed statistics about address space resource consumption.
- The top resource consumers in terms of CPU, I/O and Storage can now be identified at a glance
- The report can serve as an excellent starting point for further drill-down and analysis
- Job related GQSCAN activities have been invisible in the past. With the new report detailed statistics with regard to GQSCAN usage can now be obtained

RMF z/OS 2.2 introduces new Monitor III Sysplex reports to monitor sysplex-wide z/OS Distributed File system (zFS) usage and performance.

z/OSMF Resource Monitoring is the strategic state-of-the-art frontend for RMF Monitor III performance data. Quite recently the following features have been added:
- Collect and display historical performance data on individual metric group level
- Export the data contained in the dashboards in CSV format for later usage in a spreadsheet

Last not Least the zEvent Mobile Application:
- This mobile App is currently under development and all details and capabilities – including the application name – are subject to change
- In a nutshell the app provides the following two main features to system administrators and performance analysts:
- Receive push messages based on critical system events instantly
- Access to the RMF Data Portal and z/OSMF Resource Monitoring anywhere and every time

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   **OA45833**
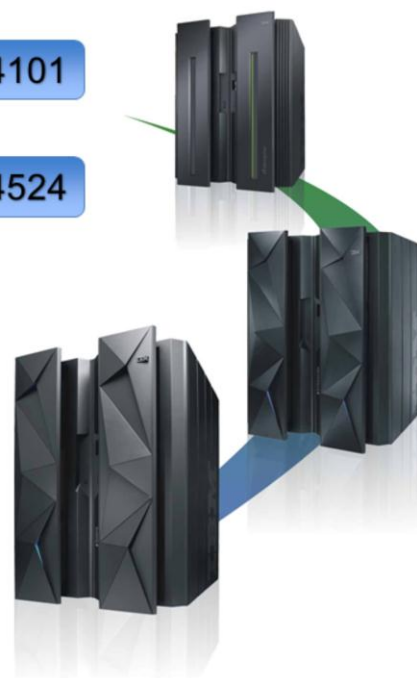
With various new function APARs, RMF exploits the new functionality of the IBM z13:

- 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
- OA45833:   z13 toleration for z/OS 1.12 and z/OS 1.13
             .

IBM®

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

Single Threaded

Multi Threaded

60

45
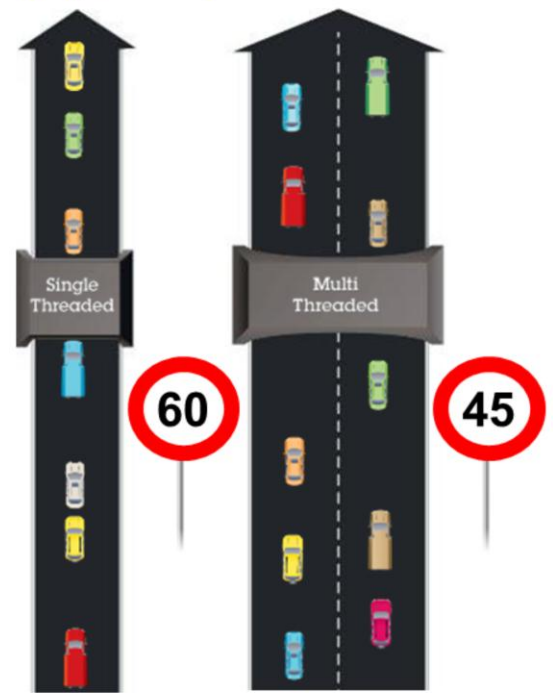
Two lanes process more traffic overall

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.

Basic idea of the SMT approach: better exploitation of processor resources with out cloning all peripheral components of a physical CPU, e.g. Cache, Register(?)

The SMT mode might not be suited for single threaded workloads (typically batch workloads).

Highway example: More overall thruput on a physical core, but speed on the individual lane is lower:

- As always, administration and switching of two lanes causes overhead
- Threads must share the peripheral processor resources, queueing and waiting cycles can occur
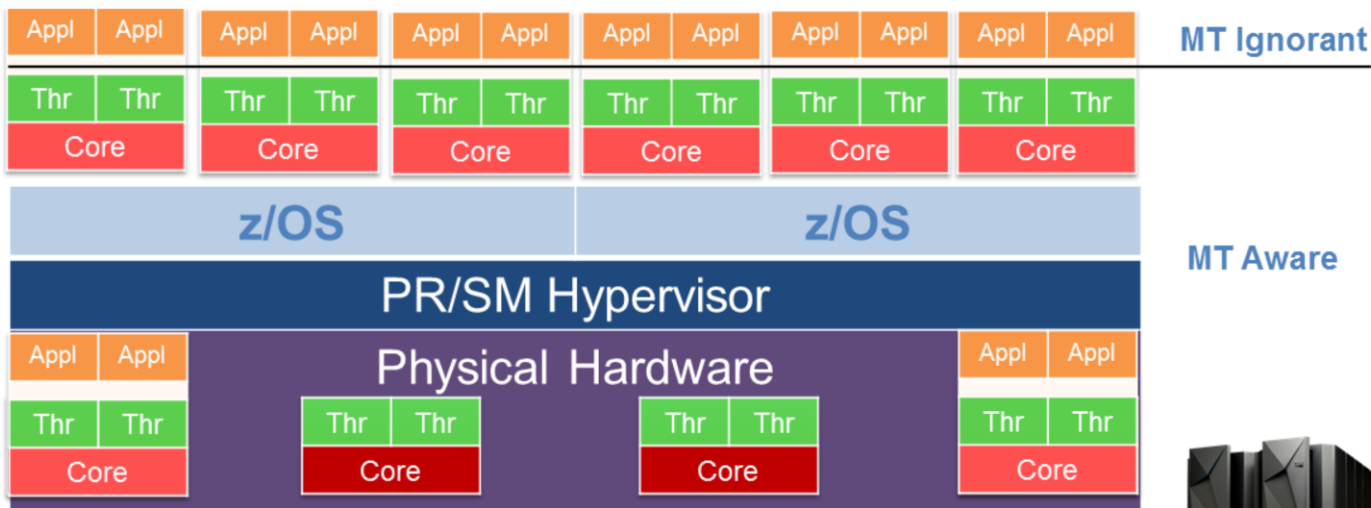
Alltogether, since with SMT more logical processing units are available, SMT helps to slow down the demand for processor resources.

Potential migration problem from zEC12 to z13: in case of single threaded workloads it is not recommended to turn on MT-2 mode.

But the overall thruput in MT-1 mode on z13 is always better than on zEC12 even though the processor speed has decreased.

This is achieved by other architectural optimizations, e.g. cache etc.

z13 – SMT Exploitation

| Appl | Appl | Appl | Appl | Appl | Appl | Appl | Appl | Appl | Appl | Appl | Appl | **MT Ignorant** |

| Thr | Thr | Thr | Thr | Thr | Thr | Thr | Thr | Thr | Thr | Thr | Thr |
| Core | Core | Core | Core | Core | Core |

z/OS   **z/OS**

PR/SM Hypervisor   **MT Aware**

Physical Hardware

Appl	Appl					Appl	Appl
Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr
Core	Core	Core	Core

- 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
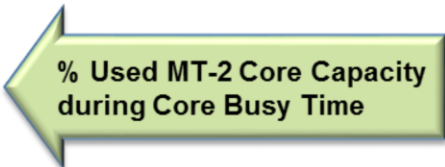
z13 - z/OS SMT Metrics

- Capacity Factor (CF)
  - ► How much work core <u>actually completes</u> 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

Actual MT-2 Efficiency

- Maximum Capacity Factor (mCF)
  - ► How much work a core <u>can complete</u> for a given workload mix at most

Estimated max MT-2 Efficiency

- 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**.

% Used MT-2 Core Capacity during Core Busy Time

- Core Utilization
  - ► Capacity in use relative to capacity total over some time interval
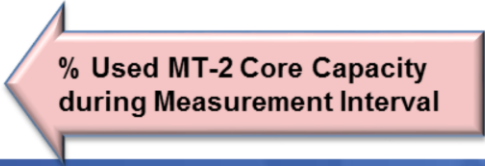 - ► Calculated as **Core Busy Time** x **Productivity**

% 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.

Is there any possibility to control the efficiency of the SMT mode?

Can I keep track of the actual speed on the 2 lanes compared to the possible maximum speed?

- The capacity factor (CF) is calculated from the number of executed instructions on both threads and the actual number of processing cycles that have been executed on the core. Usually the CF in MT-2 mode should be higher than 1.0 (except in special cases of single threaded workloads)
- The maximum capacity factor (mCF) can be achieved in case both threads are working together with maximum efficiency and no waiting cycles at all. Both, CF and mCF are subject of change: while the CF can change rapidly from interval to interval, the mCF changes are occuring slowly over time
- Core Busy Time is a metric on core level. The logical core must be dispatched on a physical core and at least one thread must be in execution state
- Average Thread Density represents the average number of executing threads during Core Busy Time
- Productivity indicates how well the available core resources have been exploited. It is calculated from the relation of CF and mCF during Core Busy Time
- Core Utilzation indicates how well my workload is suited for MT-2 mode. Since Productivity can never reach 100% the value is always below the Core Busy Time

IBM®

z13 – RMF and SMT

OA44101

- RMF enhanced with new metrics to monitor MT-2 efficiency and core utilization
- Re-interpret 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

Appl	Appl	Appl	Appl	**MT Ignorant**
Thr	Thr	Thr	Thr	**MT Aware**
Core		Core		

z/OS
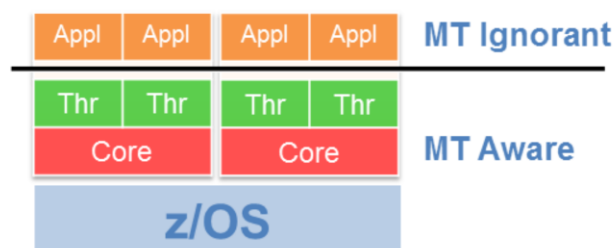
The RMF support for SMT provides new SMT related metrics to allow capacity planning and performance analysis in SMT environments.

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 the 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.

How can RMF help to monitor MT-modes efficiently:

- A couple of new MT related metrics are introduced
- Some existing metrics must be reinterpreted

First of all we need to distinguish between CPU metrics on core granularity and metrics on thread granularity.

A comprehensive description can be found in the updated RMF V2R1 documentation.

In general, RMF does NOT report the measured wallclock times, but so called MT-1 equivalent times.

Hence, it is guaranteed that the reported CPU times and service units can be compared between the z13 systems and its predecessors.

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)

The PROCVIEW CPU/CORE setting determines the content of the CPU report.

In case of CORE, new columns on CPU resp. thread level have been added.

The efficiency of MT-2 mode for zIIP engines can be monitored by means of a new MULTI-THREADING ANALYSIS section.

Remark: Since Parking is executed on core level, thread level reporting doesn'n make sense. However, the values are extracted from local MVS control blocks for the home partition.

z13 – SMT: Monitor III CPC Report

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                                                  IIP 1308.346
REM ENC  0.00                                            ABSRPTN   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:

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

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.

The WLMGL Report does not contain any new MT mode related metrics. It is a well suited example for the reinterpretation of existing metrics.

SMT in general has no effect on pricing or chargebacks:
- 4h MSU related measurements are not affected at all
- Pricing for CPU consumption on zIIP engines is not MSU related

Resource Measurement Facility

z13 – SMT: SMF Record 70-1
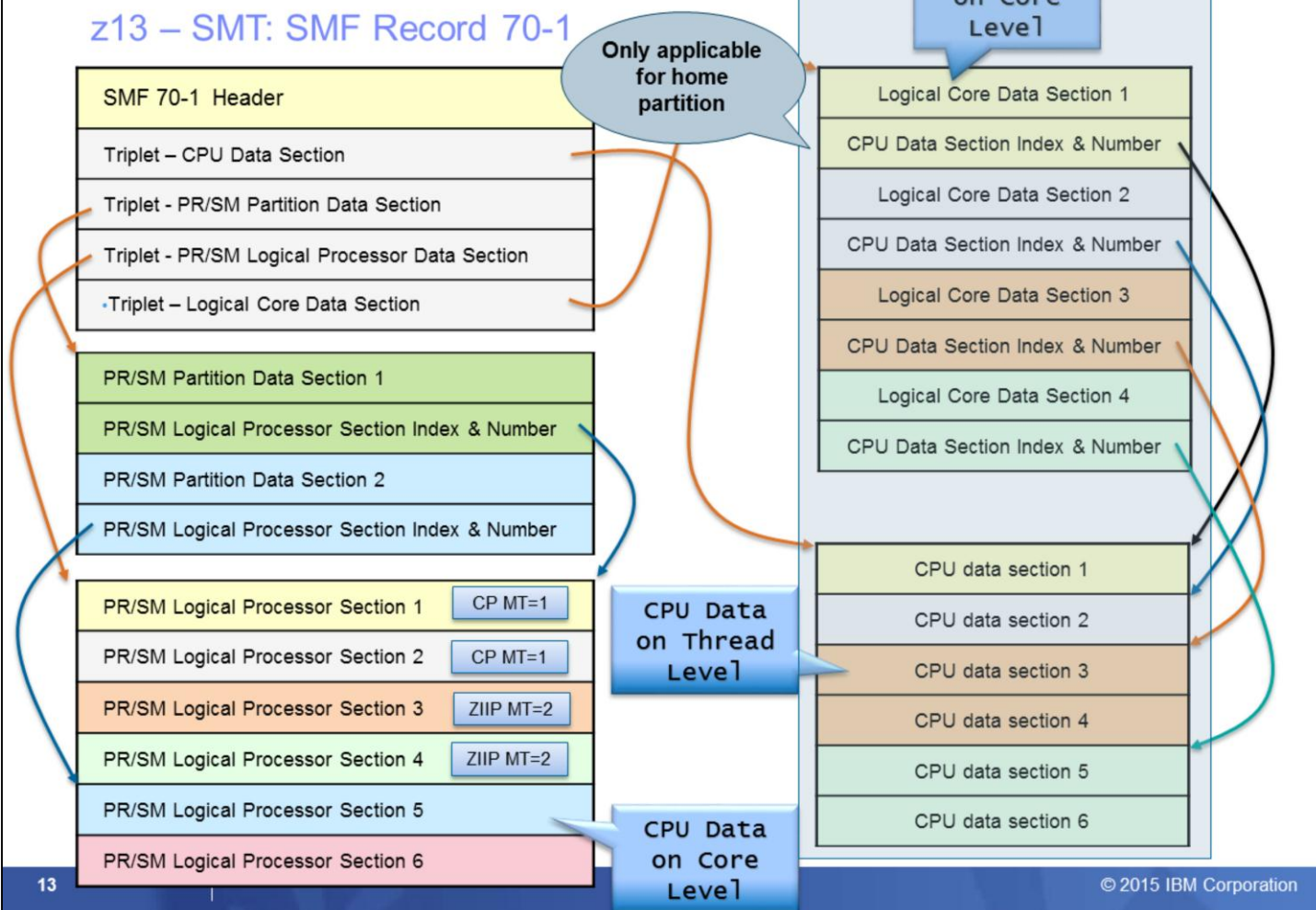
- The RMF SMT support enhances the SMF 70-1 record with SMT related fields.
- **CPU Control Data Section:**
    - Processor Class Maximum Capacity Factor Metrics:
        - SMF70MCF (CP), SMF70MCFS (zIIP)
    - Processor Class Capacity Factor Metrics:
        - SMF70CF (CP), SMF70CFS (zIIP)
    - Processor 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.

IBM®

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 - Support of Crypto Express5S Card

OA43493
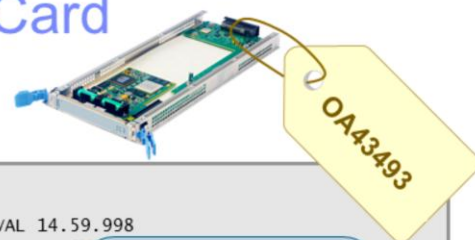
```
                        C R Y P T O   H A R D W A R E   A C T I V I T Y
z/OS V2R2              SYSTEM ID TRX2           DATE 02/01/2014         INTERVAL 14.59.998
                              RPT VERSION V2R2 RMF      TIME 11.00.00              CY
-------- CRYPTOGRAPHIC CCA COPROCESSOR ----
           -------- TOTAL -------- KEY-GEN
TYPE   ID  RATE  EXEC TIME  UTIL%    RATE
CEX5C  0   2.15   227.8     48.9     2.15
```

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

```
-------- 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
```

> **Additional ICSF service statistics**

```
-------- 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
```
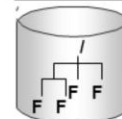
- RMF new function APAR OA43493 introduces the RMF support of the Crypto Express5S feature on the z13
- In detail, RMF collects and reports performance measurements for operations executed on Crypto Express5S CCA coprocessors (CEX5C) , PKCS11 coprocessors (CEX5P) and accelerators (CEX5A).
- The crypto 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 z13 Support**
  - ▶ RMF Statistics for Simultaneous Multithreading (SMT)
  - ▶ Extended ICSF Measurements for Crypto Express5S
  - ▶ Support for LPARs with up to 4TB Real Storage

- **SCM I/O Adapter Performance Reporting**
  - ▶ Statistics for EADM Subchannel Activities
  - ▶ Performance and Thruput on SCM Card Level

- **PCIE Activity Reporting**
  - ▶ z/OS V2.1 Postprocessor PCIE reporting
  - ▶ Additional Measurements for PCIe attached RoCE and zEDC Devices
  - ▶ z/OS V2.2 Monitor III PCIE Report

- **z/OS V2.2 Monitor III Job Resource Consumption Reporting**
  - ▶ CPU, I/O and Storage Consumption Data at a Glance
  - ▶ Detailed Statistics for Job related GQSCAN Activities

- **z/OS V2.2 zFS Reporting Enhancements**
  - ▶ Sysplex-wide Statistics for zFS Usage and Performance
  - ▶ Improved zFS Data Gathering Performance
  - ▶ Additional zFS Statistics for Shared File System environments

- **z/OSMF Resource Monitoring**
  - ▶ Historical Reporting & Spreadsheet Export

- **The zEvent Mobile Application**
  - ▶ Receive Push Messages based on critical System Events
  - ▶ Access to RMF Data Portal and z/OSMF Resource Monitoring

Storage Class Memory – aka Flash Memory – is a new tier within the memory hierarchy of the zSeries family.

RMF provides detailed usage statistics for Storage Class Memory related operations by means of a new Monitor III SCM Activity report.

IBM®

z/OS 2.2 - SCM I/O Adapter Performance Reporting

- New tier within the memory hierarchy of the zSeries family
- Also denoted as Flash Memory
- Delivers fast Solid State Drive (SSD) technology
- 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
- Assign Flash Memory to partitions like Main Memory
  - ► Flash memory allocation panel on the SE
  - ► Amount of memory initially online to a partition
  - ► Can be adjusted dynamically per command
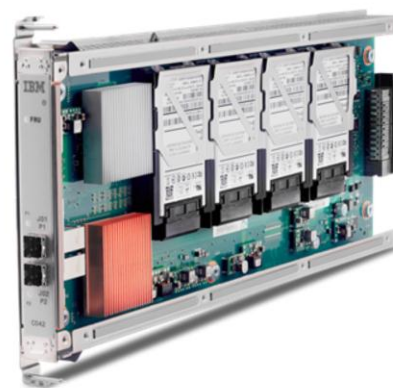
- Storage Class Memory (SCM) is a new type of memory that can be classified between main storage and auxilliary storage.
- It can be associated with a certain kind of Solid State Drive (very fast) and it is also denoted as *Flash Memory*
- Caution: The abbreviation SCM is also used for Single Chip Module ☺

z/OS 2.2 - SCM I/O Adapter Performance

- Designed for improved paging performance
- Compelling addition to traditional auxiliary 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

Access time

Layer	Access time
CPU	
Cache	< 20 ns
Random Access Memory (RAM)	< 200 ns
Flash Express	5-20 micro sec.
Solid State Drive (SSD) Storage	1-3 ms
Spinning Disk Drive	< 10 ms
WORM, Tape Library	seconds
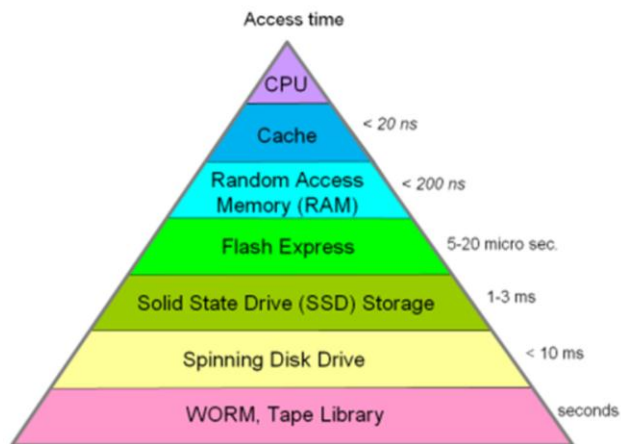
- Flash Memory is exploited by the z/OS Paging Subsystem. Hence it basically provides an extension of the auxilliary storage with an excellent performance.
- Significant benefit can be achieved when a high amount of paging data is affected. This applies especially for Pageable Large Pages of 1 MB Pages.
- High virtual memory is primarily exploited by DB2 and Java workloads. Hence, installations running those workloads should consider to benefit from Flash Express.
- Beyond high performance paging Flash Express can also supply accelerated SVC and standalone dump processing
- 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.
- The Flash Express technology is way faster then SSD technology. It provides access times within the microsecond range.

EADM Subchannels

- I/O is managed by Extended Asynchronous Data Mover subchannels (EADM)
- Similar to standard I/O subchannels, but no channel path or device number assigned
- Created automatically at IPL time
- Not tied to a particular device, no association between subchannel and SCM card pair
- Any EADM subchannel can be used to access on any SCM card assigned to the LPAR

- When z/OS needs to read or write data from Storage Class Memory, it creates a new type of channel program and issues a SSCH to a special subchannel called an Extended Asynchronous Data Mover (EADM) subchannel. EADM subchannels are similar to I/O subchannels where you can issue I/O instructions such as SSCH to run channel programs and receive I/O interrupts. However, they do not have channel paths or device numbers assigned, and they are not defined in the I/O configuration. They are created automatically at IPL time.

- Unlike I/O subchannels, which are tied to a particular device, there is no association between an EADM subchannel and a storage increment or SCM card pair. Any EADM subchannel can be used to access storage on any SCM card assigned to the LPAR. The EADM subchannel is simply a certain kind of vehicle for accessing Storage Class Memory.

z/OS 2.2 - SCM Data Gathering

- RMF Monitor III Data Gatherer collects SCM performance statistics frequently

- SCM performance data is written to the new SMF 74 subtype 10 record

- New Monitor III data gatherer option SCM | NOSCM

```
/*******************************************************************/
/* NAME:        ERBRMF04                                          */
/* DESCRIPTION: PARMLIB MEMBER WITH RMF MONITOR III GATHERER OPTIONS */
/*              (ALL OPTIONS ARE SET TO DEFAULTS)                 */
/*******************************************************************/
  CYCLE(1000)              /* SAMPLE EVERY SECOND (1000 MSEC)     */
  DATASET(STOP)            /* NO DATASET SUPPORT                  */
  DATASET(NOSWITCH)        /* APPEND TO LAST NON-FULL DATASET     */
  DATASET(WHOLD(7))        /* CONTROLS BUFFER PAGES IN STORAGE    */
  MINTIME(100)             /* LENGTH OF MINTIME                   */
NOOPTIONS                  /* DO NOT DISPLAY OPTIONS              */
  RESOURCE(*JES2,JES2)     /* SPECIFIES JES STARTED TASK NAME     */
NOSTOP                     /* RUN UNTIL OPERATOR ISSUES STOP      */
  SYNC(00)                 /* MINTIME SYNCHRONIZATION             */
  ...                      /*                                     */
  PCIE                     /* ACTIVATE PCIE DATA GATHERING        */
  SCM                      /* ACTIVATE SCM DATA GATHERING         */
  ZFS                      /* ZFS DATA GATHERING                  */
```

> SCM Data collected by default

- Data gathering option SCM / NOSCM was added to RMF Monitor III that can be used to control data collection SCM related activities
- Default value SCM is set in shipped PARMLIB member ERBRMF04
- SCM performance data is written to the new SMF 74 subtype 10 record
- Currently no RMF Postprocessor Report is provided for the SMF 74.10 record
- Third party tools can be used for after-the-facts analysis of SMF 70.10 data (e.g. MXG, EPV Technologies, SMF4U)

z/OS 2.2 - Monitor III SCM Activity Report

```
              RMF V2R2    SCM Activity

Samples: 60        System: CB8A  Date: 04/18/15  Time: 10.56.00

---------------- EADM Device/Subchannel Summary ----------------
SSCH Total      SSCH Rate      PEND Time      IOPQ Time      ICMR Time
      58          0.48           0.003          0.000          0.000

Card ID    Util(%)      Read(B/s)      Write(B/s)     Req Rate     Resp Time    IOPQ
          Part Total   Part Total     Part Total     Part Total   Part Total   Time

0000-0180 0.00  0.00   0.00  0.00    0.00  0.00     0.00  0.00   0.000 0.000   0.000
0000-01AC 0.00  0.00   0.00  0.00    0.00  0.00     0.00  0.00   0.000 0.000   0.000
0001-0100 0.00  0.00   0.00 409.6   853.3 1331     0.21  0.42   0.143 0.504   0.000
0001-012C 0.00  0.00   0.00  0.00   1126  1980     0.27  0.48   0.136 0.150   0.000
```

CPC & LPAR View

Two lines per SCM Resource

The Monitor III SCM Activity report consists of two sections:

- Header section with global EADM subchannel statistics
- Tabular section with performance and thruput statistics on SCM Card level

The EADM subchannel activities are balanced by the system and no tuning is not appliccable for this kind of resource.

Hence, the report provides just the accumulted statistics for all subchannels, no statistics on subchannel level.

One SCM Card which is plugged into the PCIe I/O drawer is also denoted as SCM resource.

Since one SCM resource can hold two internal cards, the tabular part of the SCM Activity report shows two lines per SCM resource.

In other words, the term *Card ID* consists of a prefix which identifies the SCM resource and a suffix which identifies the internal card.

The card mirrors are not visible in any way on the SCM Activity report.

The SCM card statistics are available for the local partitions as well as for the entire CPC.

- The statistics of the Monitor III SCM ISPF Report are available also by means of the Monitor III Data Portal without limitations.
- The SCM Report can be selected from the report list and basically all report columns can be displayed in the browser window.

z/OS 2.2 - RMF Distributed Data Server: SCM Metrics

Sysplex
- **MVS Image**
  - **I/O Subsystem**
    - :
    - **SCM**
      - SCM_Card
  - **Processor**
  - **Storage**
  - **Enqueue**
  - **Operator**
  - **Subsystems**
- **CPC**
  - **LPAR**
- **Coupling Facility**
  - **CF Structure**

Resource	New metric
SCM	► # of SSCH ► function pending time ► initial command response time ► IOP queue time ► SSCH rate
SCM (by card) **SCM_CARD**	► % partition utilization ► % total utilization ► average partition response time ► average total response time ► average total IOP queue time ► partition byte read rate ► partition byte write rate ► partition request rate ► total byte read rate ► total byte write rate ► total request rate

The RMF DDS Resource Model represents a composition of resources that can exist in a Parallel Sysplex environment
- MVS_IMAGE is a child resource of resource SYSPLEX
- I/O_SUBSYSTEM is a child resource of resource MVS_IMAGE
- Child resource SCM is added to resource I/O_SUBSYSTEM
- Child resource SCM_CARD is added to resource SCM

A variety of metric values that are related to resource types SCM and SCM_CARD can be requested from the RMF Distributed Data Server (DDS)

Alternatively, the browser based version of the report can be requested from the RMF Distributed Data Server (DDS) by using the following URL: http://hostname:8803/gpm/rmfm3.xml?report=SCM&resource=,sysname,MVS_IMAGE

- SMF74.10 data can be visualized with the new Postprocessor SCM Activity Report .
- This report is also available with the RMF Data Portal. The SCM Report can be selected from the report list and all report data can be displayed in the browser window.
- Name-value pairs and columns are the same as with the Monitor III SCM report.

RMF Enhancements at a Glance

- **IBM z13 Support**
  - ► RMF Statistics for Simultaneous Multithreading (SMT)
  - ► Extended ICSF Measurements for Crypto Express5S
  - ► Support for LPARs with up to 4TB Real Storage

- **SCM I/O Adapter Performance Reporting**
  - ► Statistics for EADM Subchannel Activities
  - ► Performance and Thruput on SCM Card Level

- **PCIE Activity Reporting**
  - ► z/OS V2.1 Postprocessor PCIE reporting
  - ► Additional Measurements for PCIe attached RoCE and zEDC Devices
  - ► z/OS V2.2 Monitor III PCIE Report

- **z/OS V2.2 Monitor III Job Resource Consumption Reporting**
  - ► CPU, I/O and Storage Consumption Data at a Glance
  - ► Detailed Statistics for Job related GQSCAN Activities

- **z/OS V2.2 zFS Reporting Enhancements**
  - ► Sysplex-wide Statistics for zFS Usage and Performance
  - ► Improved zFS Data Gathering Performance
  - ► Additional zFS Statistics for Shared File System environments

- **z/OSMF Resource Monitoring**
  - ► Historical Reporting & Spreadsheet Export

- **The zEvent Mobile Application**
  - ► Receive Push Messages based on critical System Events
  - ► Access to RMF Data Portal and z/OSMF Resource Monitoring

Together with the z EC12 server family two new types of PCIe cards have been introduced:

- The new z Enterprise Data Compression (zEDC) Express offering provides low-cost data compression for z/OS system services and applications
- 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 (similar than Hipersockets for LPAR to LPAR Communication)
- The existing Postprocessor PCIe Report has been extended with additional measurements
- With z/OS V2R2 RMF a complete new Monitor III PCIe Report allows to keep track of PCIe related operations in online mode

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, zBC12 and z13.
- 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.

IBM zEnterprise Data Compression capability and zEDC Express

The growth of data that needs to be captured, transferred, and stored for large periods of time is not relenting. On the contrary, software-implemented compression algorithms are costly in terms of processor resources, and storage costs are not negligible either.

The zEDC Express, an optional feature exclusive to z13, zEC12 and zBC12, addresses that problem by providing hardware-based acceleration for data compression and decompression.

The z/OS V2R1 *IBM zEnterprise* Data Compression acceleration capability uses *the zEDC Express* feature. It delivers an integrated solution to help reduce CPU consumption, optimize performance of compression-related tasks, and enable more efficient use of storage resources, while providing a lower cost of computing and also helping to optimize the cross-platform exchange of data.

IBM.

zEC12 - Shared Memory Communication

- Shared Memory Communication via Remote Direct Memory Access (SMC-RDMA or SMC-R) is a zEC12 and z13 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)

✓ **Common PCIe form factor**
✓ **Common API for retrieval of activity statistics**

10GbE RoCE Express

The 10 Gigabit Ethernet (10GbE) RoCE Express feature uses Remote Direct Access Memory over Converged Ethernet, and can provide fast memory-to-memory communications between two CPCs. It is transparent to applications.

Use of the 10GbE RoCE Express feature might help reduce consumption of CPU resources for applications that use the TCP/IP stack (such as IBM WebSphere® accessing a DB2 database). It might also help reduce network latency with memory-to-memory transfers by using Shared Memory Communications - Remote Direct Memory Access (SMC-R) in z/OS V2R1.

HiperSockets

The *HiperSockets* function is an integrated function of the CPC that provides users with attachments to up to 32 high-speed virtual local area networks with minimal system and network overhead.

HiperSockets is a function of the virtualization Licensed Internal Code (LIC) and provides LAN connectivity across multiple system images on the same CPC by performing memory-to-memory data transfers in a secure way. The HiperSockets function eliminates having to use I/O subsystem operations and having to traverse an external network connection to communicate between logical partitions in the same CPC. Therefore, HiperSockets offers significant value in server consolidation by connecting virtual servers and simplifying the Enterprise network.

Improved functions of HiperSockets (also available on zEC12, zBC12, z196, and z114) include the ability to integrate in the *intraensemble data network* (IEDN), and support for bridging to z/VM virtual switches.

z/OS - 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
```

Additional info: Obtain PCIe Information via API

IQPINFO – Obtain PCIe Information

– The IQPINFO service provides PCIe related information, including any performance statistics

– The service is described in

MVS Programming: Authorized Assembler Services Reference

– The response data area of the IQPINFO service is mapped by the macros

IQPYPERF PCIE Performance Data Return Area

IQPYPFMB PCIE Function Measurement Block

z/OS V2.1 - 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

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

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

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

z/OS V2.1 - RMF Postprocessor PCIE Activity Report

Resource Measurement Facility

▼ 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

- The General PCIE Activity section displays common hardware related statistics independent of the special purpose or type of card, e.g. Number of Load and Store operations.
- The Hardware Accelerator Activity section shows statistics which are common for all accelerator types of PCIE cards, e.g. Execution and queueing related metrics.
- Since the RoCE card is not an accelerator card, no statistics are applicable for this card type.
- The last section displays statistics related to the individual personality of the accelerator card. Hence, for the zEDC card we see compression related statistics in this section.
- Other personalities and hence additional personality related sections are subject of future extensions, e.g. a sort accelerator card might be a future candidate.

z13 - RMF PCIE Enhancements

OA44524
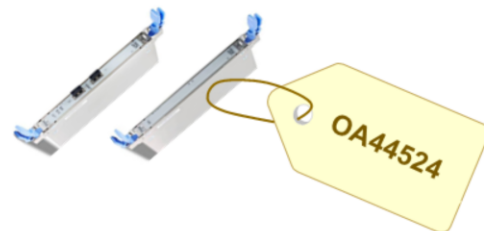
- z13 introduces new PCIE performance measurements for RDMA-over-converged-ethernet (RoCE Express) and zEnterprise data compression (zEDC) devices
  - Existing DMA read/write measurements are replaced by new PCIE function type specific measurements
  - For RoCE Express devices, there are four new measurements
    - **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.

- Due to this architectural change with z13, PCIe Read and Write statistics can now be obtained independent of the DMA address space.
- The PCIe related thruput is now reported in terms of Byte Transfer Rates and additionally as Packet Rates.
- Since on z13 the actual number of processed work units as well as the theoretical maximum number of work units can be retrieved, the PCIe adapter utilization can be calculated and reported as well.

z/OS V2.2 – RMF Monitor III PCIE Activity Report

- RMF Monitor III Data Gatherer collects PCIe performance statistics frequently

- z/OS V2.1 introduced RMF Postprocessor PCIE Activity Report for after-the-facts analysis with SMF 74.9 data

- The new RMF Monitor III PCIE Activity Report provides detailed short-term information about PCIE Express based functions. Currently supported functions are:
  - ► z Enterprise Data Compression (zEDC)
  - ► Shared Memory Communication via RDMA (SMC-R)

- New Monitor III data gatherer option PCIE | NOPCIE

V2.2

```
/**********************************************************************/
/* NAME:        ERBRMF04                                              */
/* DESCRIPTION: PARMLIB MEMBER WITH RMF MONITOR III GATHERER OPTIONS  */
/*              (ALL OPTIONS ARE SET TO DEFAULTS)                     */
/**********************************************************************/
  CYCLE(1000)             /* SAMPLE EVERY SECOND (1000 MSEC)     */
  DATASET(STOP)           /* NO DATASET SUPPORT                  */
  DATASET(NOSWITCH)       /* APPEND TO LAST NON-FULL DATASET     */
  DATASET(WHOLD(7))       /* CONTROLS BUFFER PAGES IN STORAGE    */
  MINTIME(100)            /* LENGTH OF MINTIME                   */
NOOPTIONS                 /* DO NOT DISPLAY OPTIONS              */
  RESOURCE(*JES2,JES2)    /* SPECIFIES JES STARTED TASK NAME     */
NOSTOP                    /* RUN UNTIL OPERATOR ISSUES STOP      */
  SYNC(00)                /* MINTIME SYNCHRONIZATION             */
.........................................
  OPD                     /* ACTIVATE OMVS PROCESS DATA GATHERING */
  PCIE                    /* ACTIVATE PCIE DATA GATHERING         */
  ZFS                     /* ACTIVATE ZFS DATA GATHERING          */
```

Default value PCIE set in shipped PARMLIB member ERBRMF04

- New Monitor III PCIE Activity report allows short-term performance analysis of RoCE devices and zEDC hardware accelerators. Before z/OS 2.2, performance problems on PCIe devices and hardware accelerators could only be analyzed after the facts using SMF 74.9 / RMF Postprocessor, now RMF online monitoring can be used to identify performance problems on short notice when they appear.

- Users can control whether or not they want Monitor III to collect PCIE activity data by specifying data gathering option: PCIE | NOPCIE

- Default value PCIE is set in shipped PARMLIB member ERBRMF04

z/OS V2.2 – RMF Monitor III PCIE Activity Report

- Request the Monitor III PCIE Activity Report by selection 3 from the Primary Menu & selection 14 from the Resource Report Selection Menu or enter command: *PCIE* or *PCI*
- Metrics on the report main panel are independent of the type of the exploited hardware feature and reflect the activity of the z/OS system on which RMF data collection took place.

ISPF Report

```
                        RMF V2R2   PCIE Activity                    Line 1 of 3
    Command ===>                                           Scroll ===> CSR

    Samples: 300      System: TA2   Date: 02/18/15  Time: 22.20.00  Range: 300   Sec

    -------------- Function -----------   Alloc - PCI Operations Rate - -Xfer Rate -
     ID  PCID Type Jobname  ASID Status   Time%  Load Store Block  Refr  Read Write

    006C 0204 zEDC FPGHWAM  0013 Alloc     100      0   102     0  14.0
    007C 025C zEDC FPGHWAM  0013 Alloc     100      0   102     0  14.0
    00A2 013C RoCE VTAM390  002C Alloc     100 0.113  5999     0  1.69 0.267   288
```

The Monitor III PCIE Activity Report can be used to investigate performance problems that are related to PCI Express based functions.

On the main panel, metrics are displayed that are independent of the type of the exploited hardware feature and reflect the activity of the z/OS system on which RMF data collection took place.

z/OS V2.2 – RMF Monitor III PCIE Activity Report

- Additional metrics are displayed for **zEDC Accelerators** on pop-up panel HW Accelerator And Compression Activity

```
                    RMF V2R2   PCIE Activity                    Line 1 of 3
 Command ===>                                              Scroll ===> CSR

 Samples: 300     System: TA2   Date:
 ---------------Function-----------           RMF Hardware Accelerator And Compression Activity
  ID  PCID Type Jobname  ASID Status    Press Enter to return to the Report panel.

 006C 0204 zEDC FPGHWAM  0013 Alloc                                            More:    +
 007C 025C zEDC FPGHWAM  0013 Alloc    Function ID :  006C       Alloc Time % :   99.7
 00A2 013C RoCE VTAM390  002C Alloc    Allocated   :  22.03.22   on  02/18/15

                                        Hardware Accelerator
                                         Time Busy %    : 0.286        Transfer Rate :   4.87

                                        Request
                                         Execution Time :   28.0       Std. Deviation:     8.07
                                         Queue Time   :     65.7       Std. Deviation:      140
                                         Size         :     47.6

          DELAYJ                         Buffer Pool
                                          Memory Size  :     16        Utilization   :       0

                                                       Compression   Decompression
 • No additional metrics,                Request Rate :    102          0.437
   no pop-up panel for                   Throughput   :    2.91        0.009
   RoCE devices                          Ratio        :    2.79        0.652
```

If cursor is placed on one of the cursor sensitive fields *Function ID, Function PID, or Function Type*, additional metrics are displayed for **zEDC Accelerators** on pop-up panel HW Accelerator And Compression Activity

If cursor is placed on one of the cursor sensitive fields *Function ID, Function PID, or Function Type*, message "No additional information available" is displayed when selected PCIE function is a **RoCE devic**e

Alternatively, the browser based version of the report can be requested from the RMF Distributed Data Server (DDS)

z/OS V2.2 – RMF Data Portal PCIE Activity Report

The statistics of the Monitor III PCIE ISPF Report are available also by means of the Monitor III Data Portal without limitations.

The PCIE Report can be selected from the report list and basically all report columns can be displayed in the browser window.

But due to the high number of report columns of you need to use the slider in order to see the rightmost colums as well.

As an alternative, you can also switch to the vertical report view to see all measurements for a selected device at a glance.

z/OS V2.2 – RMF Distributed Data Server: PCIE Metrics

Sysplex

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

Resource	New metric
PCIE	► … by PCIE function
PCIE_Function	► % allocation time ► % buffer pool utilization ► % time busy ► buffer pool memory size ► compression ratio ► compression request rate ► compression throughput ► decompression ratio ► decompression request rate ► decompression throughput ► received packets rate ► request execution time ► request execution time standard deviation ► request queue time ► request queue time standard deviation ► request size ► transfer rate ► transfer read rate ► transfer write rate ► transmitted packets rate ► work unit rate ► PCI adapter utilization ► PCI load operations rate ► PCI refresh operations rate ► PCI store block operations rate ► PCI store operations rate

RMF Distributed Data Server (DDS) supports new resource types *PCIE* and *PCIE Function* to allow performance analysis of RoCE devices and zEDC hardware accelerators by DDS API exploiters

RMF DDS Resource Model represents a composition of resources that can exist in a Parallel Sysplex environment
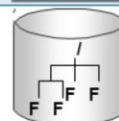- MVS_IMAGE is a child resource of resource SYSPLEX
- I/O_SUBSYSTEM is a child resource of resource MVS_IMAGE
- Child resource PCIE is added to resource I/O_SUBSYSTEM
- Child resource PCIE_FUNCTION is added to resource PCIE

A variety of metric values that are related to resource types PCIE and PCIE_FUNCTION can be requested from the RMF Distributed Data Server (DDS)

The metrics are available as single valued metrics for the PCIE_FUNCTION resource and as list valued metrics for the PCIE resource.
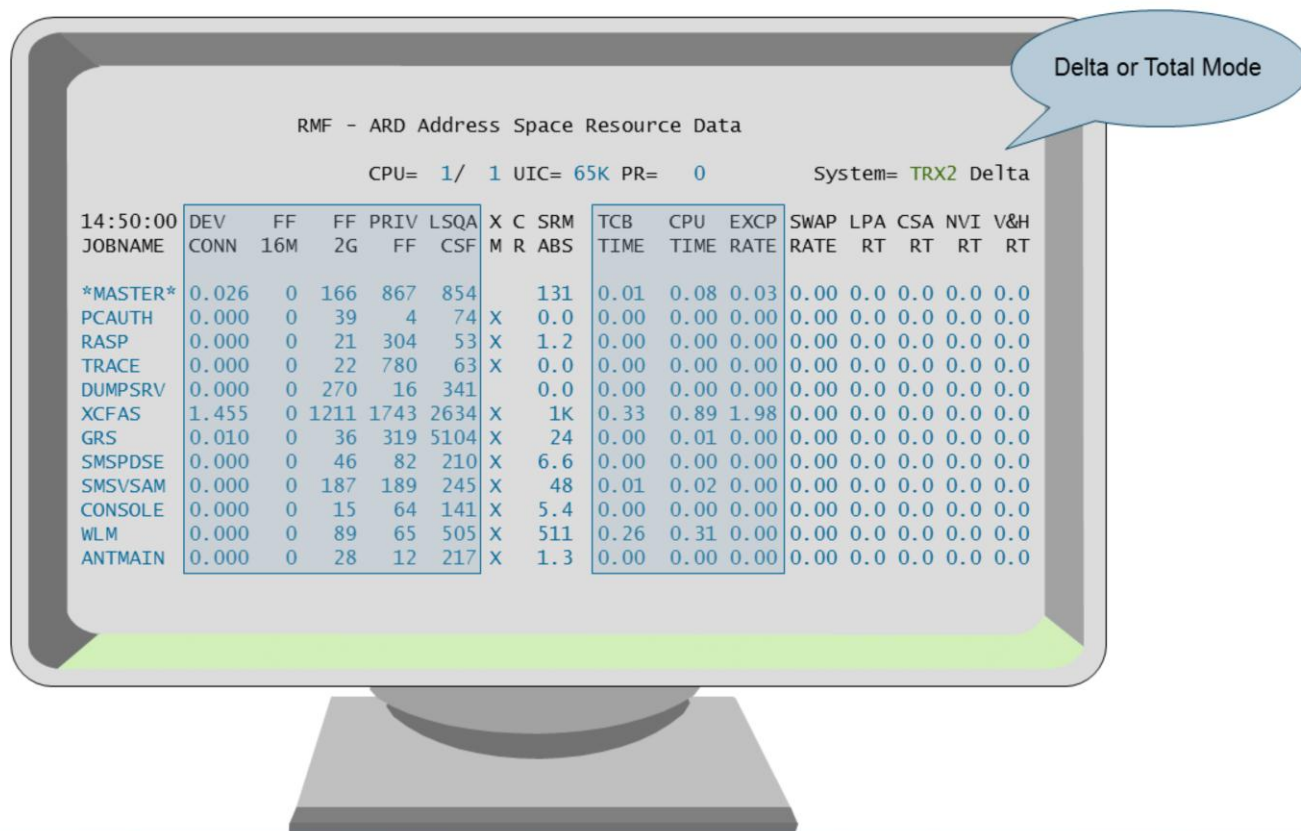
RMF Enhancements at a Glance

- **IBM z13 Support**
  - ► RMF Statistics for Simultaneous Multithreading (SMT)
  - ► Extended ICSF Measurements for Crypto Express5S
  - ► Support for LPARs with up to 4TB Real Storage

- **SCM I/O Adapter Performance Reporting**
  - ► Statistics for EADM Subchannel Activities
  - ► Performance and Thruput on SCM Card Level

- **PCIE Activity Reporting**
  - ► z/OS V2.1 Postprocessor PCIE reporting
  - ► Additional Measurements for PCIe attached RoCE and zEDC Devices
  - ► z/OS V2.2 Monitor III PCIE Report

- **z/OS V2.2 Monitor III Job Resource Consumption Reporting**
  - ► CPU, I/O and Storage Consumption Data at a Glance
  - ► Detailed Statistics for Job related GQSCAN Activities

- **z/OS V2.2 zFS Reporting Enhancements**
  - ► Sysplex-wide Statistics for zFS Usage and Performance
  - ► Improved zFS Data Gathering Performance
  - ► Additional zFS Statistics for Shared File System environments

- **z/OSMF Resource Monitoring**
  - ► Historical Reporting & Spreadsheet Export

- **The zEvent Mobile Application**
  - ► Receive Push Messages based on critical System Events
  - ► Access to RMF Data Portal and z/OSMF Resource Monitoring

The Monitor III Job Usage Report complements the Monitor III reporting suite with detailed statistics about address space resource consumption.

- The top resource consumers in terms of CPU, I/O and Storage can now be identified at a glance
- The report can serve as an excellent starting point for further drill-down and analysis
- Job related GQSCAN activities have been invisible in the past. With the new report detailed statistics with regard to GQSCAN usage can now be obtained

Job Resource Consumption Reporting: Monitor II ARD Report

Delta or Total Mode

```
              RMF - ARD Address Space Resource Data

              CPU=  1/  1 UIC= 65K PR=    0          System= TRX2 Delta

14:50:00 DEV   FF   FF PRIV LSQA X C SRM  TCB   CPU  EXCP SWAP LPA CSA NVI V&H
JOBNAME  CONN  16M   2G   FF  CSF M R ABS  TIME  TIME RATE RATE  RT  RT  RT  RT

*MASTER* 0.026   0  166  867  854       131  0.01  0.08 0.03 0.00 0.0 0.0 0.0 0.0
PCAUTH   0.000   0   39    4   74 X     0.0  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
RASP     0.000   0   21  304   53 X     1.2  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
TRACE    0.000   0   22  780   63 X     0.0  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
DUMPSRV  0.000   0  270   16  341 X     0.0  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
XCFAS    1.455   0 1211 1743 2634 X      1K  0.33  0.89 1.98 0.00 0.0 0.0 0.0 0.0
GRS      0.010   0   36  319 5104 X      24  0.00  0.01 0.00 0.00 0.0 0.0 0.0 0.0
SMSPDSE  0.000   0   46   82  210 X     6.6  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
SMSVSAM  0.000   0  187  189  245 X      48  0.01  0.02 0.00 0.00 0.0 0.0 0.0 0.0
CONSOLE  0.000   0   15   64  141 X     5.4  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
WLM      0.000   0   89   65  505 X     511  0.26  0.31 0.00 0.00 0.0 0.0 0.0 0.0
ANTMAIN  0.000   0   28   12  217 X     1.3  0.00  0.00 0.00 0.00 0.0 0.0 0.0 0.0
```

RMF Monitor II provides job related resource consumption by means of the following reports

- Address Space Resource Data
- Address Space State Data
- Address Space SRM Data

The data is reported on snapshot base for a certain interval (Delta mode) or it is accumulated for the entire life cycle of the address space (Total mode)

However, there is now comprehensive report for job related resource consumption, the data is spread on three different reports.
For this reason, with z/OS V2R2 RMF the new Monitor III Job Usage Report is introduced.

The ARD report gives information on the system resources that are used by an address space. The information provided in this report includes, for example, information on processor time, paging, and central storage.

The ARD report enables you to determine which jobs are creating performance problems.

From the Monitor II ARD Report, the following statistics are now also available in Monitor III

- Device Connect Time
- Fixed Frame Counts on Virtual Storage location level
- TCB Times and Total CPU Times
- EXCP Rates

Job Resource Consumption Reporting: Monitor II ASD Report

```
          RMF - ARD Address Space State Data

          CPU=  1/ 1 UIC= 65K PR=   0          System= TRX2 Delta

14:50:00          S C  R  DP  CS    CS  TAR  X  PIN  TX  SWAP   WSM
JOBNAME  SRVCLASS P L  LS PR   F   TAR  WSS  M   RT  SC   RV     RV

*MASTER* SYSTEM   1 NS     FF 6353        0       0   0    0
PCAUTH   SYSTEM   1 NS     FF  129        0  X    0   0    0
RASP     SYSTEM   1 NS     FF  363        0  X    0   0    0
TRACE    SYSTEM   1 NS     FF  872        0  X    0   0    0
DUMPSRV  SYSTEM   1 NS     FF  583        0       0   0    0
XCFAS    SYSTEM   1 NS     FF 7814        0  X    0   0    0
GRS      SYSTEM   1 NS     FF 14.1K       0  X    0   0    0
SMSPDSE  SYSTEM   1 NS     FF 4320        0  X    0   0    0
SMSVSAM  SYSTEM   1 NS     FF 3310        0  X    0   0    0
CONSOLE  SYSTEM   1 NS     FF 4477        0  X    0   0    0
WLM      SYSTEM   1 NS     FF 29.0K       0  X    0   0    0
ANTMAIN  SYSTEM   1 NS     FF 1511        0  X    0   1   998
```

The ASD report gives an overview of the current state of an address space. Basically, the report tells you where each address space is and what it is doing.

You can use the ASD report, for example, to determine which jobs are using large amounts of central storage or which jobs are being swapped excessively and why the swapping is occurring.

If you have a workload delaying your application, you can check the workloads dispatching priority (**DP PR**) on the ASD report, and change it if necessary.

From the Monitor II ASD Report, the following statistics are now also available in Monitor III

- Service Class and Service Class Period
- Dispatching Priority
- Central Storage Frame Counts

Job Resource Consumption Reporting: Monitor II ASRM Report

```
                RMF  -  ASRM Address Space SRM Data

              CPU=  1/  1 UIC= 65K PR=   0            System= TRX2 Delta

14:33:58          S TRANS    TRANS      TX  TX     TX      TX      TX      TX    SESS
JOBNAME  SRVCLASS P ACTIVE   CUR RES    CT  SC     CPU     MSO     IOC     SRB   TOTAL

*MASTER* SYSTEM   1  31:10 31:10:51    1   0   5183K      0  51347 17549K 22784K
PCAUTH   SYSTEM   1  31:10 31:10:54    1   0     42       0      5    158    205
RASP     SYSTEM   1  31:10 31:10:54    1   0   20855      0      0 172162 193017
TRACE    SYSTEM   1  31:10 31:10:54    1   0    374       0     80    357    811
DUMPSRV  SYSTEM   1  31:10 31:10:36    1   0   3296       0   2910   1711   7917
XCFAS    SYSTEM   1  31:10 31:10:41    1   0  87313K      0  1181K 142.2M 230.7M
GRS      SYSTEM   1  31:10 31:10:54    1   0  17289K      0    180  3351K 20640K
SMSPDSE  SYSTEM   1  31:10 31:10:54    1   0   1111K      0      0 172760  1283K
SMSVSAM  SYSTEM   1  31:10 31:10:41    1   0   3157K      0   2945   4832K  7992K
CONSOLE  SYSTEM   1  31:10 31:10:54    1   0   3670K      0   2815 897165  4570K
WLM      SYSTEM   1  31:10 31:10:41    1   0  70296K      0    910 13016K 83312K
ANTMAIN  SYSTEM   1  31:10 31:10:41    1   1 217423      0   3360  49190 269973
ANTAS000 SYSSTC   1  31:10 31:10:36    1   1  21204      0   2945   2657  26806
```

The ASRM report gives an overview of the system resources that are used by an address space.

The report gives, for example, information on processor service, storage service, and I/O service.

The report enables you to determine which jobs are using which services and whether certain jobs are creating performance problems by making excessive use of system services.

From the Monitor II ASRM Report, the following statistics are now also available in Monitor III
- Service Class and Service Class Period
- Transaction Active Times
- Transaction Current Resident Times
- Transaction Counts

z/OS V2.2 – Monitor III USAGE Report

- **Request the Monitor III USAGE Report by selection 1 from the Primary Menu & selection 4A from the Overview Report Selection Menu or enter command:**
  USAGE [job_class, service_class] or *USG*

> Identify Top Resource Consumers at a Glance

```
                          RMF V2R2    Job Oriented Usage

 Samples: 60       System: TRX1  Date: 04/18/15  Time: 10.56.00  Range: 60  Sec


              Service   --- I/O ---   --- CPU ---   - Storage -   ----- QScan ----
 Jobname   CX Class     Conn   EXCP   Total   TCB   Total Fixed   Total Resct Time

 XCFAS     S  SYSTEM    0.446  1.97   0.25   0.11    7754  2384      0   0.0      0
 BHBE      T  TSOCLASS  0.103  3.90   0.07   0.07      21   623      1   0.0   2581
 *MASTER*  S  SYSTEM    0.042  0.00   0.02   0.00    6323  1107      0   0.0      0
 SMF       S  SYSTEM    0.028  0.00   0.00   0.00     900   210      0   0.0      0
 CATALOG   S  SYSTEM    0.027  0.17   0.03   0.03    1824   228      0   0.0      0
 GRS       S  SYSTEM    0.020  0.00   0.01   0.01   14136   451      0   0.0      0
 JES2      S  SYSSTC    0.010  0.38   0.03   0.02    9277  1041      0   0.0      0
 NET       S  SYSSTC    0.010  0.00   0.01   0.00    3050   138      0   0.0      0
 DFSZFS    S  SYSSTC    0.008  0.60   0.00   0.00   30660   499      0   0.0      0
 OMVS      S  SYSTEM    0.006  0.17   0.00   0.00   16098   356      0   0.0      0
 SMS       S  SYSSTC    0.004  0.93   0.00   0.00     548    89      0   0.0      0
 PAGENT    SO SYSSTC    0.003  9.45   0.01   0.01    2978 18072      0   0.0      0
 HZSPROC   SO SYSSTC    0.000  0.00   0.00   0.00    5125   183      0   0.0      0
```

The new Monitor III Job Usage Report provides a comprehensive overview about job related resource consumption.

The report is aranged accordingly to the primary resource categories:

- I/O
- CPU
- Storage

By default the report is sorted by descending I/O Connect Time

In the ISPF Version of the report the counts in the visible colums are related to the current Monitor III Range.

Furthermore the total accumulated counts (as applicable by means of the Monitor II Total Mode) are also stored in the Monitor III ISPF table.

Hence, those counts can also be made visible by using the Monitor III Format Table Definition Utility ERBRMFU.

In addition to the general resource consumption statistics, the new report displays statistics about job related GQSCAN activities:

From the API perspective, GQSCAN requests can be initiated by GQSAN macro or ISGQUERY REQINFO=QSCAN. (ISGQUERY REQINFO=QSCAN is the IBM recommended replacement of GQSCAN)

The following counts are reported:

- QScan Total: Total requests is the accummulated number of QScan requests for the address space
- QScan Resct: Resct is the average number of resources returned by Qscan requests for the address space
- Qscan Time: Time is the average QScan request time in microseconds for the address space

z/OS V2.2 – Monitor III USAGE Report: Cursor Sensitivity Targets

```
                    RMF V2R2    Job Oriented Usage

    Samples: 60       System: TRX1  Date: 04/18/15  Time: 10.56.00  Range: 60  Sec


              Service  --- I/O ---  --- CPU ---  - Storage -  ----- QScan ----
    Jobname   CX Class  Conn    EXCP  Total   TCB  Total Fixed  Total Resct Time

    XCFAS     S  SYSTEM  0.446   1.97  0.25  0.11   7754  2384      0   0.0      0
    *MASTER*  S  SYSTEM  0.042   0.00  0.02  0.00   6323  1107      0   0.0      0
    SMF       S  SYSTEM  0.028   0.00  0.00  0.00    900   210      0   0.0      0
    CATALOG   S  SYSTEM  0.027   0.17  0.03  0.03   1824   228      0   0.0      0
    GRS       S  SYSTEM  0.020   0.00  0.01  0.01  14136   451      0   0.0      0
    JES2      S  SYSSTC  0.010   0.38  0.03  0.02   9277  1041      0   0.0      0
    NET       S  SYSSTC  0.010   0.00  0.01  0.00   3050   138      0   0.0      0
    DFSZFS    S  SYSSTC  0.008   0.60  0.00  0.00  30660   499      0   0.0      0
    OMVS      S  SYSTEM  0.006   0.17  0.00  0.00  16098   356      0   0.0      0
    SMS       S  SYSSTC  0.004   0.93  0.00  0.00    548    89      0   0.0      0
    PAGENT    SO SYSSTC  0.003   9.45  0.01  0.01   2978 18072      0   0.0      0
    HZSPROC   SO SYSSTC  0.000   0.00  0.00  0.00   5125   183      0   0.0      0
```

DELAYJ	Filter	DEV	PROCU	STORF

The Latest and Greatest   11/6/2015   © 2015 IBM Corporation

In case of general resource contention, the Job Usage report can serve as an excellent starting point for further analysis and problem determination.

In this context the Monitor III Cursor Sensitivity feature can be used to drill-down problems efficiently.

z/OS V2.2 – RMF Monitor III USAGE Report

- The set of report options common to Monitor III address space related reports is applicable for the USAGE report
- Use RO command to filter by address space type or WLM service class

```
                     RMF Delay Report Options: USAGE              Line 1 of 2
Command ===>                                            Scroll ===> CSR

Change or verify parameters. To exit press END.
Changes will apply to DELAY, DEV, ENQ, HSM, JES, PROC, PROCU, STOR, STORC,
STORF, STORM, USAGE and XCF.

  Class            ===> ALL        Classes: ALL TSO BATCH STC ASCH OMVS
  Service class ===> *ALL          *ALL or one of available service classes below

  Jobs             ===> NO         View job selection/exclusion panel next (YES NO)

                          Available Service classes
BTCHDEF      GPMSERVE    OE          OMVSKERN    STCDEF      TSODEF      SYSTEM
SYSSTC
```

YES

```
                     RMF Delay Report Options: USAGE             Line 1 of 18
Command ===>                                            Scroll ===> CSR

Select (S), exclude (X), or fill-in jobs for report. Press END.

Sel   Jobname      Sel   Jobname      Sel   Jobname      Sel   Jobname      Sel   Jobname
-     _____           _____           _____           _____           _____

S     *ALL               *MASTER*           ALLOCAS            ANTAS000           ANTMAIN
      APPC               ASCH               ASCHINT            AUXMON             AXR
      AXR04              BJAG               BKGE               BKGEDDSF           BLUCIUS
      BMAI               BPMU               BPXOINIT           CATALOG            CAZO
      CBDQDISP           CEA                CFZCIM             CONSOLE            CSF
```

The Latest and Greatest   11/6/2015   © 2015 IBM Corporation

The set of common report options for all Monitor III address space related reports will be applicable for the new USAGE report as well.

The RO command displays the Report Options panel which allows you to filter the report by address space type or WLM service class.

Once you specify *YES* in the *Jobs* input field, the Job Selection/Exclusion panel is displayed in order to filter the report by specific job names.

In addition, you can invoke the report directly by specifying an address space type or service class name as parameter (e.g. *USAGE S* or *USAGE SYSSTC*).

The new Monitor III USAGE report will be added to the standard reports which are available in XML format by means of the RMF Distributed Data Server.

Therefore all statistics of the Monitor III Job Usage ISPF Report are available also by means of the Monitor III Data Portal without limitations.

The USAGE Report can be selected from the report list and basically all report columns can be displayed in the browser window, also by using the following URL:

http://hostname:8803/gpm/rmfm3.xml?report=USAGE&resource=,sysname,MVS_IMAGE

Sorting is possible in the RMF Data Portal: Just one click on the column header brings the job with the highest consumption count to the top.

But due to the high number of report columns you need to use the slider in order to see the rightmost colums as well.

So, as an alternative, you can also switch to the vertical report view to see all measurements for a selected device at a glance.

Clicking on a jobname will display all the values for this job in one single view.

IBM®

z/OS V2.2 – RMF USAGE Report: RMF Distributed Data Server Metrics

Sysplex

MVS Image
- I/O Subsystem
- Processor
- Storage
- Enqueue
- Operator
- Subsystems

CPC
- LPAR

Coupling Facility
- CF Structure

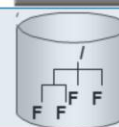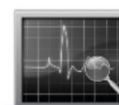| Resource	New metric
MVS_IMAGE	▶ transaction active time by job ▶ transaction resident time by job ▶ # transactions by job ▶ # QSCAN generic requests by job ▶ # QSCAN specific requests by job ▶ # QSCAN resources by job ▶ # QSCAN resources standard deviation by job ▶ QSCAN request time by job ▶ QSCAN request time standard deviation by job
STORAGE	▶ # frames fixed below 16 MB by job ▶ # frames fixed above 16 MB by job ▶ # frames fixed above 2 GB by job
I/O_SUBSYSTEM	▶ total device connect time by job ▶ connect time by job ▶ EXCP rate by job
PROCESSOR	▶ total CPU time by job ▶ CPU time by job ▶ total TCB time by job ▶ TCB time by job

DDS metrics extracted from the Job USAGE Report

In addition a subset of the metrics contained in the USAGE report were added to the job related DDS metrics. Hence, these metrics are supported by the RMF Performance Data Portal and the z/OSMF Resource Monitoring GUI.

Since the Job Usage Report contains comprehensive address space related statistics, the new DDS metrics which are extracted from this report are spread across multiple resources.

RMF Enhancements at a Glance

- IBM z13 Support
  - ▶ RMF Statistics for Simultaneous Multithreading (SMT)
  - ▶ Extended ICSF Measurements for Crypto Express5S
  - ▶ Support for LPARs with up to 4TB Real Storage

- SCM I/O Adapter Performance Reporting
  - ▶ Statistics for EADM Subchannel Activities
  - ▶ Performance and Thruput on SCM Card Level

- PCIE Activity Reporting
  - ▶ z/OS V2.1 Postprocessor PCIE reporting
  - ▶ Additional Measurements for PCIe attached RoCE and zEDC Devices
  - ▶ z/OS V2.2 Monitor III PCIE Report

- z/OS V2.2 Monitor III Job Resource Consumption Reporting
  - ▶ CPU, I/O and Storage Consumption Data at a Glance
  - ▶ Detailed Statistics for Job related GQSCAN Activities

- z/OS V2.2 zFS Reporting Enhancements
  - ▶ Sysplex-wide Statistics for zFS Usage and Performance
  - ▶ Improved zFS Data Gathering Performance
  - ▶ Additional zFS Statistics for Shared File System environments

- z/OSMF Resource Monitoring
  - ▶ Historical Reporting & Spreadsheet Export

- The zEvent Mobile Application
  - ▶ Receive Push Messages based on critical System Events
  - ▶ Access to RMF Data Portal and z/OSMF Resource Monitoring

RMF z/OS 2.2 introduces new Monitor III Sysplex reports to monitor sysplex-wide z/OS Distributed File system (zFS) usage and performance.

z/OS 2.2 - zFS Reporting Enhancements
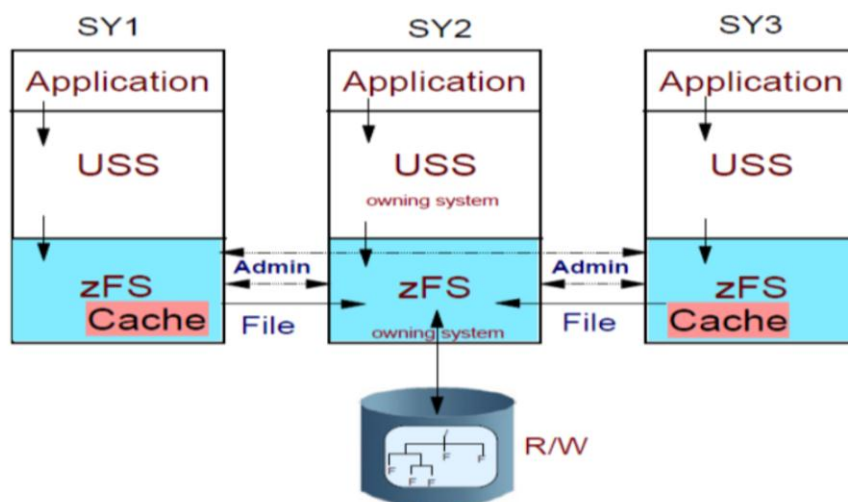
Rationale:

- Existing RMF Monitor III single system reports ZFSSUM and ZFSACT offered no possibility to monitor details of zFS related to sysplex awareness of zFS file systems.
- Some customers observed performance problems when gathering zFS performance data.

The existing RMF Monitor III zFS reports (introduced with z/OS 1.7) has some disadvantages:

- Since the reports are single system reports, it is not possible to monitor zFS activity in shared file system environments:
    - There is no sysplex wide view of zFS activity
    - There are no zFS statistics related to sysplex awareness of zFS file systems
- The current Monitor III data gatherer uses multiple zFS interface calls to collect the zFS data. In case of system environments with a larger number of zFS aggregates and file systems, performance problems were observed that caused delays in overall RMF Monitor III data gathering

RMF for z/OS 2.2 introduces new Monitor III zFS sysplex reports that allow to monitor shared file system environments effectively. The Monitor III zFS data gatherer uses a new zFS interface that reduces the number of zFS interface calls so that RMF zFS data gathering performance is improved.

z/OS 2.2 - zFS Reporting Enhancements

- z/OS V2.2 introduces three new Monitor III zFS reports:
  - zFS Overview report
  - zFS File System report
  - zFS Kernel report

```
                        RMF Sysplex Report Selection Menu
Selection ===>

Enter selection number or command for desired report.

Sysplex Reports
        1 SYSSUM    Sysplex performance summary           (SUM)
        2 SYSRTD    Response time distribution            (RTD)
        3 SYSWKM    Work Manager delays                   (WKM)
        4 SYSENQ    Sysplex-wide Enqueue delays           (ES)
        5 CFOVER    Coupling Facility overview            (CO)
        6 CFSYS     Coupling Facility systems             (CS)
        7 CFACT     Coupling Facility activity            (CA)
        8 CACHSUM   Cache summary                         (CAS)
        9 CACHDET   Cache detail                          (CAD)
       10 RLSSC     VSAM RLS activity by storage class    (RLS)
       11 RLSDS     VSAM RLS activity by data set         (RLD)
       12 RLSLRU    VSAM LRU overview                     (RLL)
       13 ZFSOVW    zFS Overview                          (ZFO)
       14 ZFSFS     zFS File System                       (ZFF)
       15 ZFSKN     zFS Kernel                            (ZFK)
```

- Sysplex-wide data on:
  - zFS response time / wait times
  - zFS XCF activity
  - zFS Cache activity
  - zFS activity / capacity by File System

- Data helps to control the zFS environment according to
  - Monitoring of zFS activity in Shared File System environments
  - Tuning of Cache sizes
  - Monitoring of File System Performance:
    - Discover Bottlenecks and
    - Balance File System I/O
  - Capacity Control for File Systems

RMF for z/OS V2.2 introduces three new RMF Monitor III zFS reports:

- zFS Overview report
- ZFS File System report
- zFS Kernel report

The new reports are sysplex reports.

They can be invoked from the RMF Sysplex Report Selection Menu or by following commands:

- **ZFSOVW** or **ZFO**:   zFS Overview report
- **ZFSFS** or **ZFF** :   zFS File System report
- **ZFSKN** or **ZFK** :   zFS Kernel report

Gathering of zFS activity data is controlled by new Monitor III gatherer option:

NOZFS | ZFS

With z/OS V2.2 the zFS data gathering default is changed back from NOZFS (z/OS 2.1) to ZFS.

z/OS 2.2 – zFS Overview report

- Monitor III zFS Overview report (ZFSOVW) provides a summary of zFS activity, wait percentages and cache statistics on the current sysplex.
- Helps to discover bottlenecks and tune cache behaviour.

```
        RMF V2R2    zFS Overview      - UTCPLXCB          Line 1 of 3


Samples: 60   Systems: 3  Date: 04/21/15 Time: 05.02.00  Range: 60  Sec


                              ---------- Cache Activity -----------
System         -----Wait%------   ---User---     --Vnode---     -Metadata-
               I/O  Lock Sleep    Rate Hit%      Rate Hit%      Rate Hit%

CB8C           21.3  0.0  36.2    40.50  100    153.2 99.2     41.22 99.1
CB8D            0.0  0.0   9.2    3.500  100    16.05  100     19.80 99.8
CB86            6.8  0.0   0.1     5775 84.2    17238 87.7     59821 93.5
```

Fields in Monitor III zFS Overview report (ZFSOVW) :

System   Name of the system running zFS.

Wait%   The following Wait percentages are reported:
- **I/O**   Percentage of time that zFS requests had to wait for I/O completion.
- **Lock**   Percentage of time that zFS requests had to wait for locks.
- **Sleep**   Percentage of time that zFS requests had to wait for events.

Dashes (----) in these fields indicate that RMF is unable to calculate a reasonable value.

Cache Activity section

User   The user file cache is for caching regular user files that are larger than 7K. The measured statistics have the following meanings:
- **Rate**   Total number of read and write requests per second made to the user file cache.
- **Hit%**   Percentage of read and write requests to the user file cache that completed without accessing the DASDs.

Vnode   The vnode cache is used to hold virtual inodes. An inode is a data structure related to a file in the file system, holding information about the file's user and group ownership, access mode and type.

The measured statistics have the following meanings:
- **Rate**   Number of read and write requests per second made to the vnode cache.
- **Hit%**   Percentage of read and write requests to the vnode cache that completed without accessing the DASDs.

Metadata   The metadata cache is used for file system metadata and for files smaller than 7K. It resides in the primary z/FS address space. The measured statistics have the following meanings:
- **Rate**   Number of read and write requests per second made to the metadata cache.
- **Hit%**   Percentage of read and write requests to the metadata cache that completed without accessing the DASDs.

From the **zFS Overview Report**, you can navigate to a variety of detail information using cursor-sensitive control.

The following details pop-up panels can be displayed:

- **I/O Details by Type** pop-up panel from any value in the **Wait%** section.
- **User Cache Details** pop-up panel from any value in the **Cache Activity – User** section.
- **Vnode Cache Details** pop-up panel from any value in the **Cache Activity – Vnode** section
- **Metadata Cache Details** pop-up panel from any value in the **Cache Activity – Metadata** section

Fields in Monitor III ZFSOVW report I/O Details pop-up panel:

System	Name of the system running zFS.
Count	Total number of I/O requests of the indicated type.
Waits	Number of zFS requests waiting for an I/O completion of the indicated I/O type.
Cancl	Number of cancelled zFS requests during an I/O request of the indicated type, for example, a user tried to delete a file during a pending I/O to this file's metadata.
Merge	Number of merges of two I/O requests into a single request because of better performance.
Type	Type of the I/O request (I/O for metadata, log data or user file data).

z/OS 2.2 – zFS Overview Cache Details

Fields in Monitor III ZFSOVW report User Cache Details pop-up panel:

System	Name of the system running zFS.
Size	Total number of pages in the user file cache.
Total Pages	Number of zFS requests waiting for an I/O completion of the indicated I/O type.
Free Pages	Total number of free pages in the user file cache.
Segments	Total number of allocated segments in the user file cache.
Storage fixed	Shows whether the size of the user file cache storage is fixed. If the zFS parameter user_cache_size is set to 'fixed', then zFS reserves real storage for use by zFS only. The 'fixed option' helps to improve performance during data access and can be applied if you have enough real storage available.
Read Rate	Number of read requests per second made to the user file cache.
Read Hit%	Percentage of read requests to the user file cache that completed without accessing the DASD.
Read Dly%	Percentage of delayed read requests to the user file cache. A read request is delayed if it must wait for pending I/O, for example, because the file is in a pending read state due to asynchronous read ahead from DASD to the user file cache.
Async Read Rate	Number of readaheads per second.
Write Rate	Number of write requests per second made to the user file cache.
Write Hit%	Percentage of write requests to the user file cache that completed without accessing the DASD.
Write Dly%	Percentage of delayed write requests to the user file cache. The following reasons are counted as write request delays: Write wait: a write must wait for pending I/O. Write faulted: a write to a file needs to perform a read from DASD. If a write-only updates a part of a file's page, and this page is not in the user file cache, then the page must be read from DASD before the new data is written to the cache.
Scheduled Write Rate	Number of scheduled writes per second.
Read%	Percentage of read requests, based on the sum of read and write requests.
Dly%	Percentage of delayed requests.
Page Reclaim Writes	Total number of page reclaim writes. A page reclaim write action writes one segment of a file from the user file cache to DASD. Page reclaim writes are performed to reclaim space in the user file cache. If page reclaim writes occur too often in relation to the write rate, then the user file cache may be too small.
Fsyncs	Total number of requests for file synchronization (fsync) between user file cache and DASD.

Fields in Monitor III ZFSOVW report Vnode Cache Details pop-up panel:

System	Name of the system running zFS.
Size	Total size of the vnode cache.

Vnodes

Total	Number of currently allocated vnodes in the vnode cache. If more vnodes are requested than are currently available, then zFS dynamically allocates more vnodes.
Size	Size of a vnode data structure in bytes.
Ext.#	Number of extended vnodes.
Ext. Size	Size of an extended vnode data structure in bytes.
Open	Number of currently open vnodes.
Held	Number of vnodes currently held in zFS by USS.

Requests

Total	Number of requests to the vnode cache.
Rate	Number of requests per second made to the vnode cache.
Hit%	Percentage of requests to the vnode data that found the target vnode data structures in the vnode cache. High hit rates indicate a favorable zFS environment, because each miss involves initialization of vnode data structures in the vnode cache.
Alloc	Number of requests to create new vnodes (for operations such as create or mkdir).
Delete	Number of requests to delete vnodes (for operations such as remove or failed creates or mkdirs).

Fields in Monitor III ZFSOVW report Metadata Cache Details pop-up panel:

System	Name of the system running zFS.
Size	Total size of the metadata cache.
Buffers	Total number of buffers in the metadata cache. The buffer size is 8K.
Storage fixed	Shows whether the size of the metadata cache storage is fixed. If the zFS parameter meta_cache_size is set to 'fixed', then zFS reserves real storage for use by zFS only. The 'fixed option' helps to improve performance during data access and can be applied if you have enough real memory available.
Total	Number of requests made to the metadata cache.
Rate	Number of requests per second made to the metadata cache.
Hit%	Percentage of requests to the metadata cache completing without accessing the DASD.
Updates	Number of updates made to buffers in the metadata cache.
Partial writes	Number of times that only half of an 8K metadata block needed to be written.

z/OS 2.2 – zFS File System report

- Monitor III zFS File System report (ZFSFS) provides detailed measurements of zFS file system activity for single file systems in compatibility mode aggregates.
- Helps to monitor performance and capacity limits of file systems.

```
                    RMF V2R2   zFS File System  - UTCPLXCB        Line 1 of 568


Samples: 60     Systems: 3   Date: 04/21/15  Time: 05.02.00  Range: 60    Sec

------ File System Name -------------------           I/O  Resp Read  XCF
               System    Owner    Mode    Size Usg%  Rate  Time  %    Rate
APIRWW.DB2
               *ALL      CB86     RW S   1320M  1.0  0.000 0.000  0.0 0.000
NETVIEW.V6R1M0C.ZFS
               *ALL      CB86     RW S     12M 49.8  0.000 0.000  0.0 0.000
OMVS.CB8C.JAVATEST.OUTPUT.ZFS
               *ALL      CB8C     RW S   9849M  0.3  0.000 0.000  0.0 0.000
OMVS.CB8C.JAVATEST.ZFS
               *ALL      CB8C     RW S   4932M  0.7  0.133 0.005  100 0.000
OMVS.CB8D.JAVATEST.OUTPUT.ZFS
               *ALL      CB8D     RW S   9849M  3.8  0.000 0.000  0.0 0.000
OMVS.CB8D.JAVATEST.ZFS
               *ALL      CB8D     RW S     13G 79.9  0.067 0.003  100 0.000
OMVS.CB86.JAVATEST.OUTPUT.ZFS
               *ALL      CB86     RW S     12G  0.3  0.000 0.000  0.0 0.000
```

Fields in Monitor III zFS File System report (ZFSFS):

File System Name	File system name.
System	Name of the system connected to the file system. In the first data line for a file system, the name is '*ALL' to indicate that this line shows the SYSPLEX view of the data rather than a single system view.
Owner	Name of owning system.
Mode	Mount mode of the file system. Possible values are: **RW** mounted in read-write mode. **RO** mounted in read-only mode. **NM** not mounted. **QS** not available because the aggregate is quiesced. The mount mode is followed by an **S** if the file system is using zFS sysplex sharing (RWSHARE).
Size	Maximum logical size of the file system (in Bytes).
Usg%	Percentage of currently used space by the file system.
I/O Rate	The rate of read and write requests per second (directory and file) made by applications to this file system.
Resp Time	Average response time in milliseconds for read and write requests made by applications to this file system.
Read%	Percentage of read operations contained in 'I/O Rate'.
XCF Rate	The rate of read and write XCF calls per second to the server.

z/OS 2.2 – zFS File System report: Report Options

- The contents of the Monitor III ZFSFS report can be tailored by report options (use RO command on ZFSFS panel).
- Example: Detailed statistics for one file systems

```
                 RMF zFS File System Report Options        Line 1 of 284

Change or verify parameters. To exit press END.
Changes will apply to the ZFSFS report.

Name    ===> OMVS.CB8D.JAVATEST.ZFS
                  ALL or one of the available zFS file systems below
Detail  ===> YES     Show single system data (YES or NO) in ZFSFS report

                     Available zFS File Systems
OMVS.CB8C.JAVATEST.OUTPUT.ZFS
OMVS.CB8C.JAVATEST.TESTSUIT.ZFS
OMVS.CB8C.JAVATEST.ZFS
OMVS.CB8D.JAVAT
OMVS.CB8D.JAVAT
OMVS.CB8E.JAVAT
OMVS.CB8E.JAVAT
```

```
                         RMF V2R2   zFS File System  - UTCPLXCB        Line 1 of 5

Samples: 60      Systems: 3    Date: 04/21/15  Time: 05.02.00  Range: 60    Sec

------ File System Name --------------------             I/O  Resp Read  XCF
                  System    Owner    Mode    Size Usg%  Rate  Time  %    Rate
OMVS.CB8D.JAVATEST.ZFS
                  *ALL      CB8D     RW S    13G 79.9  0.067 0.003 100 0.000
                  CB8C      CB8D     RW S        0.0  0.000 0.000  0.0 0.000
                  CB8D      CB8D     RW S    13G 79.9  0.067 0.003 100 0.000
                  CB86      CB8D     RW S        0.0  0.000 0.000  0.0 0.000
```

58

The Report Options panel for the zFS File System report allows you to specify options for this report.

Fields in Monitor III zFS File System report options panel:

Name Specify either ALL or the name of one of the zFS file systems available in the sysplex as shown in the field **Available File Systems**, which provides a list of all zFS file systems that are currently defined to the sysplex. You can use an asterisk ('*') as the last character of the file system name as a wild card. When a wild card is used, all file systems whose names start with the specified character sequence before the asterisk are reported on, no matter which characters follow.

Detail Specify the desired level of detail in the zFS File System report:

   NO The report contains summary data for the sysplex only.

   YES The report contains data for the sysplex and all single systems.

If the list of file names is too long to fit on the first page, this report options panel can be scrolled up and down using function keys F7 and F8.

z/OS 2.2 – zFS File System report: Report Options

- Example: Using ALL for 'Name' and Details=YES, complete information for all file systems on all systems is provided.

```
                    RMF V2R2   zFS File System  - UTCPLXCB        Line 1 of 1420


Samples: 60      Systems: 3   Date: 04/21/15  Time: 05.02.00   Range: 60      Sec

------ File System Name --------------------             I/O  Resp Read  XCF
               System    Owner    Mode    Size Usg%     Rate  Time  %    Rate
APIRWW.DB2
               *ALL      CB86     RW S   1320M  1.0    0.000 0.000  0.0 0.000
               CB8C      CB86     RW S           0.0   0.000 0.000  0.0 0.000
               CB8D      CB86     RW S           0.0   0.000 0.000  0.0 0.000
               CB86      CB86     RW S   1320M  1.0    0.000 0.000  0.0 0.000
NETVIEW.V6R1M0C.ZFS
               *ALL      CB86     RW S     12M 49.8    0.000 0.000  0.0 0.000
               CB8C      CB86     RW S           0.0   0.000 0.000  0.0 0.000
               CB8D      CB86     RW S           0.0   0.000 0.000  0.0 0.000
               CB86      CB86     RW S     12M 49.8    0.000 0.000  0.0 0.000
OMVS.CB8D.JAVATEST.OUTPUT.ZFS
               *ALL      CB8D     RW S   9849M  3.8    0.000 0.000  0.0 0.000
               CB8C      CB8D     RW S           0.0   0.000 0.000  0.0 0.000
               CB8D      CB8D     RW S   9849M  3.8    0.000 0.000  0.0 0.000
               CB86      CB8D     RW S           0.0   0.000 0.000  0.0 0.000
OMVS.CB8D.JAVATEST.ZFS
               *ALL      CB8D     RW S     13G 79.9    0.067 0.003  100 0.000
               CB8C      CB8D     RW S           0.0   0.000 0.000  0.0 0.000
               CB8D      CB8D     RW S     13G 79.9    0.067 0.003  100 0.000
               CB86      CB8D     RW S           0.0   0.000 0.000  0.0 0.000
```

z/OS 2.2 – zFS File System report: Cursor Sensitivity Targets

```
                    RMF V2R2   zFS File System  - UTCPLXCB        Line 1 of 5

Samples: 60      Systems: 3   Date: 04/21/15  Time: 05.02.00  Range: 60    Sec

------ File System Name --------------------             I/O  Resp Read  XCF
             System    Owner    Mode    Size Usg%  Rate  Time  %    Rate
OMVS.CB8D.JAVATEST.ZFS
              *ALL      CB8D      RW S    13G 79.9  0.067 0.003 100 0.000
              CB8C      CB8D      RW S        0.0   0.000 0.000 0.0 0.000
              CB8D      CB8D      RW S    13G 79.9  0.067 0.003 100 0.000
              CB8                   zFS File System Details

        File System Name : OMVS.CB8D.JAVATEST.ZFS
        Mount
        Point : /CB8D/javatest

        System : CB8D              Owner : CB8D              Mode : RW S

        -------------- Read --------------    ------------- Write -------------
        --- Appl --- ---- XCF ----   Aggr    --- Appl --- ---- XCF ----   Aggr
         Rate   Resp   Rate   Resp   Rate     Rate   Resp   Rate   Resp   Rate
                Time          Time                   Time          Time
        0.067  0.003  0.000  0.000  0.000    0.000  0.000  0.000  0.000  204.8

        Vnodes                 :     3       USS held vnodes         :      2
        Open objects           :     0       Tokens                  :      3
        User cache 4k pages :     0          Metadata cache 8k pages :     10

        ENOSPC errors          :     0       Disk I/O error          :      0
        XCF comm. failures  :     0          Cancelled operations    :      0
```

The Late

From the zFS File System Report, you can navigate to a variety of detail information using cursor-sensitive control. If you place the cursor on any of the lines with file system values, a pop-up window appears showing the details for this file system.

z/OS 2.2 – zFS Kernel report

- Monitor III zFS Kernel report (ZFSKN) provides measurements counting the calls made to zFS from z/OS UNIX and the average response time of zFS requests.
- Helps to monitor zFS Kernel performance and determine appropriate tuning options

```
                          RMF V2R2   zFS File System  - UTCPLXCB         Line 1 of 3


Samples: 60        Systems: 3   Date: 04/21/15  Time: 05.02.00  Range: 60     Sec

System        - Request Rate -  --- XCF Rate ---  - Response Time -
Name          Local    Remote    Local    Remote   Local    Remote

CB8C          54.18    219.7     0.000    3.417    0.017    0.140
CB8D          9.150    49.52     0.000    0.150    0.149    2.996
CB86          35298    577.5     0.000    46.95    0.610    3.453
```

Fields in Monitor III zFS Kernel report (ZFSKN):

System Name	Name of the system running zFS. In the context of requests against file systems, this is the name of the requesting system.
Request Rate	Rate of zFS requests during the report interval for file systems which are locally and remotely owned. A file system is locally owned if the requesting system is also the owner of the file system. It is remotely owned if the owner of the file system is not the requesting system.
XCF Rate	Rate of zFS requests during the report interval requiring data from another system via XCF, both for locally and remotely owned file systems.
Response Time	Average time in milliseconds required for the completion of the zFS requests during the report interval for locally and remotely owned file systems.

z/OS V2.2 – RMF Data Portal ZFS Reports

- The browser based version of ZFSOVW, ZFSFS and ZFSKN can be requested from the RMF Distributed Data Server (DDS) by use of following URL (example is for ZFSKN): http://sysf.boebl.de:8803/gpm/rmfm3.xml?report=ZFSKN&resource=",,SYSPLEX"

RMF Report [,UTCPLXCB,SYSPLEX] : ZFSKN (zFS Kernel)

Time Range: 03/03/2015 08:48:00 - 03/03/2015 08:49:00

System Name	Request Rate Local	Request Rate Remote	XCF Rate Local	XCF Rate Remote	Response Time Local	Response Time Remote
CB8A	113.5	0.000	0.067	0.000	17269	0.000
CB8C	378.4	554.6	0.000	72.53	26.00	10473
CB86	26.40	1569	0.000	119.7	14.00	4302
CB89	21.80	524.0	0.000	72.42	11.00	10198

The new Monitor III ZFSOVW, ZFSFS and ZFSKN reports were added to the standard reports which are available in XML format by means of the RMF Distributed Data Server.

Therefore all statistics of the new Monitor III zFS ISPF reports are available also by means of the Monitor III Data Portal without limitations.

The zFS Reports can be selected from the list of full RMF reports that are available for the SYSPLEX resource and basically all report columns can be displayed in the browser window.

As alternative you can invoke a zFS report by use of following URLs :

ZFSOVW report: http://hostname:8803/gpm/rmfm3.xml?report=ZFSOVW&resource=,,SYSPLEX

ZFSFS report: http://hostname:8803/gpm/rmfm3.xml?report=ZFSFS&resource=,,SYSPLEX

ZFSKN report: http://hostname:8803/gpm/rmfm3.xml?report=ZFSKN&resource=,,SYSPLEX

The performance metrics for the new zFS sysplex reports were added to the zFS related DDS metrics. The metrics are associated to the Aggegate, ZFS and SYSPLEX resources of the DDS resource tree.
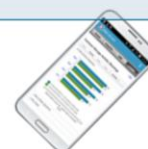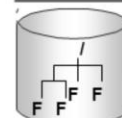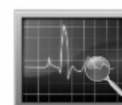
Basically the DDS is the data source for z/OSMF Resource Monitoring, the state–of–the–art graphical workstation frontend for RMF performance data.

All DDS metrics can be configured for continuous monitoring with z/OSMF in terms of metrics groups and dashboards.

Hence, lets talk about z/OSMF Resource Monitoring and the most recent enhancements...

RMF Enhancements at a Glance

- **IBM z13 Support**
  - ► RMF Statistics for Simultaneous Multithreading (SMT)
  - ► Extended ICSF Measurements for Crypto Express5S
  - ► Support for LPARs with up to 4TB Real Storage

- **SCM I/O Adapter Performance Reporting**
  - ► Statistics for EADM Subchannel Activities
  - ► Performance and Thruput on SCM Card Level

- **PCIE Activity Reporting**
  - ► z/OS V2.1 Postprocessor PCIE reporting
  - ► Additional Measurements for PCIe attached RoCE and zEDC Devices
  - ► z/OS V2.2 Monitor III PCIE Report

- **z/OS V2.2 Monitor III Job Resource Consumption Reporting**
  - ► CPU, I/O and Storage Consumption Data at a Glance
  - ► Detailed Statistics for Job related GQSCAN Activities

- **z/OS V2.2 zFS Reporting Enhancements**
  - ► Sysplex-wide Statistics for zFS Usage and Performance
  - ► Improved zFS Data Gathering Performance
  - ► Additional zFS Statistics for Shared File System environments

- **z/OSMF Resource Monitoring**
  - ► Historical Reporting & Spreadsheet Export

- **The zEvent Mobile Application**
  - ► Receive Push Messages based on critical System Events
  - ► Access to RMF Data Portal and z/OSMF Resource Monitoring

z/OSMF Resource Monitoring is the strategic state-of-the-art frontend for RMF Monitor III performance data. Quite recently the following features have been added:

- Collect and display historical performance data on individual metric group level
- Export the data contained in the dashboards in CSV format for later usage in a spreadsheet

z/OSMF 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...

	Start Dashboard		ve Frames
	Pause Dashboard		
SYSTEM.1	2907	Retrieve Historical Data...	FTPS1 [0041] ▶	564
		Stop Retrieving Historical Da...	
		Save Dashboard	CBIND1 [005D] ▶	499
SYSSTC.1	4436	Rename Dashboard...	
		Modify Settings...	
STCDEF.1 ▶	3530	Add Metric...	
	203000	Modify Metric...	
		Export...	
GPMSERVE.1 ▶	261	Delete Metric or Group	
	428	Arrange Metrics or Gro...	
		Rename Metric Group.	
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)

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

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

| The Latest and Greatest | 11/6/2015 | © 2015 IBM Corporation

z/OSMF Resource Monitoring – Spreadsheet Export

Export Dashboard

Export Dashboard

✓ Step1: Select Scope and Intervals

⇨ **Step2: Select Metrics and Resources**

Select the metrics and resources to include when exporting the metric group.

Metric group to export: Active & Fixed Frames (WLM View)

Metrics:

S1.*.STORAGE # frames fixed by WLM service class period
S1.*.STORAGE # frames active by WLM service class period

Resources:

OMVS.1
OMVS.3
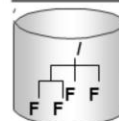SHRDB2.1
STCHI.1
STCLO.1
STCMD.1
SYSOTHER.1
SYSSTC.1
SYSTEM.1

[< Back] [Next >] [Finish] [Cancel]

⇨ **Step1: Select Scope and Intervals**

✓ Step2: Select Metrics and Resources

Select the scope a...

Scope of Data to Ex...

◉ Metric group:

    Active & Fixed Fran

    ◉ Include all metri

    ○ Include a selecti

○ Entire dashboard wi

Range of Intervals t...

Start time of first interv

03/03/2014 14:25:00

End time of last interva

03/03/2014 14:45:00

[< Back] [Next >] [Finish] [Cancel]

📊 Storage Soaker_Active & Fixed Frames (WLM View).csv

📊 Storage Soaker_Active Frames.csv

📊 Storage Soaker_Fixed Frames.csv

IBM.

RMF Enhancements at a Glance

- **IBM z13 Support**
  - ▶ RMF Statistics for Simultaneous Multithreading (SMT)
  - ▶ Extended ICSF Measurements for Crypto Express5S
  - ▶ Support for LPARs with up to 4TB Real Storage

- **SCM I/O Adapter Performance Reporting**
  - ▶ Statistics for EADM Subchannel Activities
  - ▶ Performance and Thruput on SCM Card Level

- **PCIE Activity Reporting**
  - ▶ z/OS V2.1 Postprocessor PCIE reporting
  - ▶ Additional Measurements for PCIe attached RoCE and zEDC Devices
  - ▶ z/OS V2.2 Monitor III PCIE Report

- **z/OS V2.2 Monitor III Job Resource Consumption Reporting**
  - ▶ CPU, I/O and Storage Consumption Data at a Glance
  - ▶ Detailed Statistics for Job related GQSCAN Activities

- **z/OS V2.2 zFS Reporting Enhancements**
  - ▶ Sysplex-wide Statistics for zFS Usage and Performance
  - ▶ Improved zFS Data Gathering Performance
  - ▶ Additional zFS Statistics for Shared File System environments

- **z/OSMF Resource Monitoring**
  - ▶ Historical Reporting & Spreadsheet Export

- **The zEvent Mobile Application**
  - ▶ Receive Push Messages based on critical System Events
  - ▶ Access to RMF Data Portal and z/OSMF Resource Monitoring

Last not Least the zEvent Mobile Application:

- This mobile App is currently under development and all details and capabilities – including the application name – are subject to change
- In a nutshell the app provides the following two main features to system administrators and performance analysts:
  - Receive push messages based on critical system events instantly
  - Access to the RMF Data Portal and z/OSMF Resource Monitoring anywhere and every time

z/OSMF Resource Monitoring & Mobile Devices

Very convenient to know right now that the Sysplex of my boss has got enough CSA and SQA...

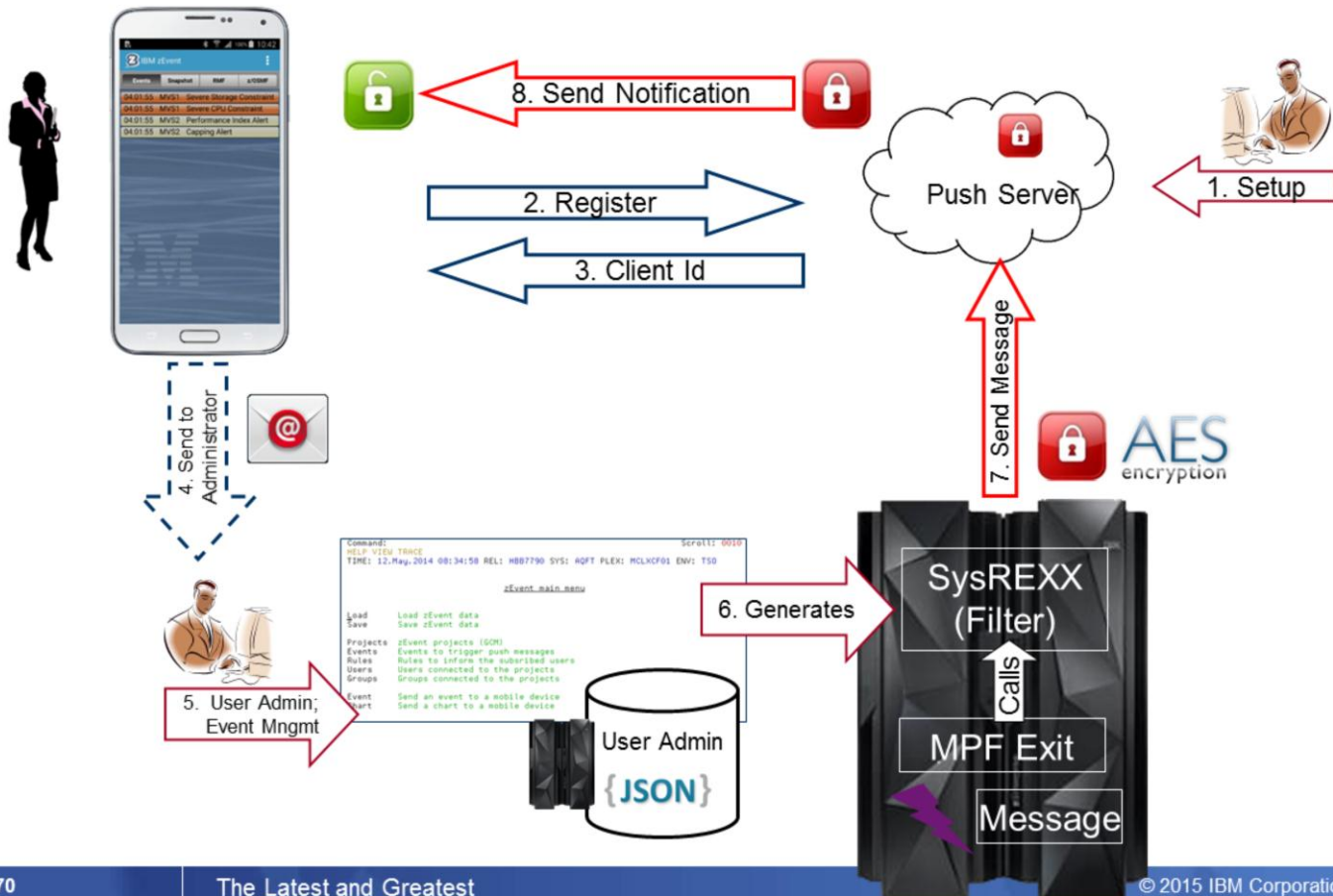Cross platform performance monitoring anytime and everywhere!

z/OSMF supports by default the recent versions of Mozilla Firefox and Microsoft Internet Explorer

However, z/OSMF can be used together with the most browsers available for mobile devices as well

Cross platform performance monitoring anytime and everywhere: Try z/OSMF on your tablet PC or smartphone!

zEvent Mobile Application

- Receive push messages based critical system events
- Access to z/OS performance data from mobile devices
- Single point of control for the enterprise
- Developed with IBM Mobile First Studio
- Prototype status (Android OS)
- Current availability target is 1Q 2016
- Supported Monitoring Facilities:
  - ▶ RMF Performance Data Portal
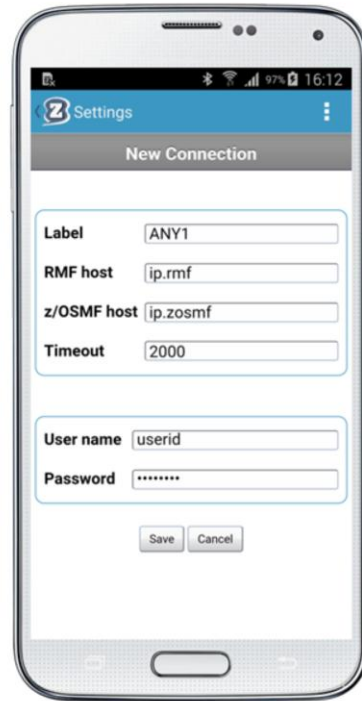 - ▶ z/OSMF Resource Monitoring

The Latest and Greatest

11/6/2015

© 2015 IBM Corporation

zEvent Mobile Application

8. Send Notification

2. Register

3. Client Id

Push Server

1. Setup

4. Send to Administrator

7. Send Message

AES encryption

5. User Admin; Event Mngmt

User Admin {JSON}

6. Generates

SysREXX (Filter)

Calls

MPF Exit

Message

The Latest and Greatest

zEvent Mobile Application

- Connect to all z/OS systems in the enterprise
- Direct connection from the App to the z/OS system or via Mobile First Server

IBM®

zEvent Mobile Application

- Monitor z/OS system performance with the RMF Data Portal
- Exploit the extended capabilities of z/OSMF Resource Monitoring

The Latest and Greatest   11/6/2015   © 2015 IBM Corporation
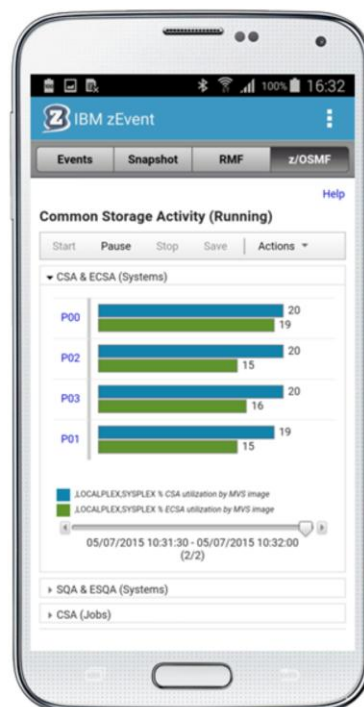
Information and Tools

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

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

- Documentation and news:
  - ▶ RMF Report Analysis, SC34-2665
  - ▶ RMF User's Guide, SC34-2664
  - ▶ RMF Programmer's Guide, SC34-2667
  - ▶ Latest version of PDF files can be downloaded from:
    www.ibm.com/systems/z/os/zos/library/bkserv/v2r2pdf/#ERB

The Latest and Greatest

11/6/2015

Function Reference

Function	Availability	
zEnterprise Data Compression Express Reporting	z/OS 2.1 RMF	**zEC12 Enhancements**
Shared Memory Communication Reporting	z/OS 2.1 RMF	
Support for Simultaneous Multithreading (SMT)	APAR OA44101	
Extended Measurements for PCIe attached Devices	APAR OA44524	
Additional ICFS Statistics for Crypto Express5S	APAR OA43493	**z13 Enhancements**
Support for LPARs with up to 4 TB real storage	APAR OA44503	
Support for z13 IBM Integrated Coupling Adapter (ICA SR)	APAR OA44502	
RMF z13 Toleration for z/OS V1R10 and V1R11	APAR OA45890	
RMF z13 Toleration for z/OS V1R12 and V1R13	APAR OA45833	
SCM I/O Adapter Performance Reporting	z/OS 2.2 RMF	
RMF Monitor III PCIE Online Reporting	z/OS 2.2 RMF	
RMF Monitor III Job Resource Consumption Reporting	z/OS 2.2 RMF	
RMF Monitor III zFS Reporting Enhancements	z/OS 2.2 RMF	
z/OSMF Resource Monitoring – Historical Reporting	APAR PI08825	
z/OSMF Resource Monitoring – Spreadsheet Export	APAR PI08825	
zEvent Mobile Application	1Q 2016 (Subject to Change)	