

# IBM Education Assistance for z/OS V2R2

Item: RMF Monitor III PCIE Support

Element/Component: RMF



## Agenda

- Presentation Objectives
- Overview
- Usage & Invocation: RMF Monitor III Data Gatherer Option
- Usage & Invocation: RMF Monitor III PCIE Activity Report
- Usage & Invocation: RMF DDS PCIE Activity Support
- Presentation Summary
- Appendix



## Presentation Objectives

- Explain purpose/usage of RMF z/OS V2R2 Monitor III PCIE Support
- Data gathering option PCIE / NOPCIE was added to RMF Monitor III that can be used to control data collection of hardware features
  - RDMA (Remote Direct Memory Access) support for z/OS over Converged Enhanced Ethernet (RoCE)
  - IBM zEnterprise Data Compression (zEDC) capability using zEDC Express
- New Monitor III PCIE Activity report allows short-term performance analysis of RoCE devices and zEDC hardware accelerators
- 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



## Overview

- Problem Statement / Need Addressed
  - Performance problems on RoCE devices and zEDC hardware accelerators can only be analyzed after the facts using SMF 74.9 / RMF Postprocessor
- Solution
  - The new RMF Monitor III PCIE Activity Report and the Distributed Data Server API allow short-term performance analysis of PCIe devices and hardware accelerators
- Benefit / Value
  - RMF online monitoring can be used to identify performance problems related to PCIe devices and hardware accelerators



## Usage & Invocation: RMF Monitor III Data Gatherer Option

- 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

```

/*****
/* 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

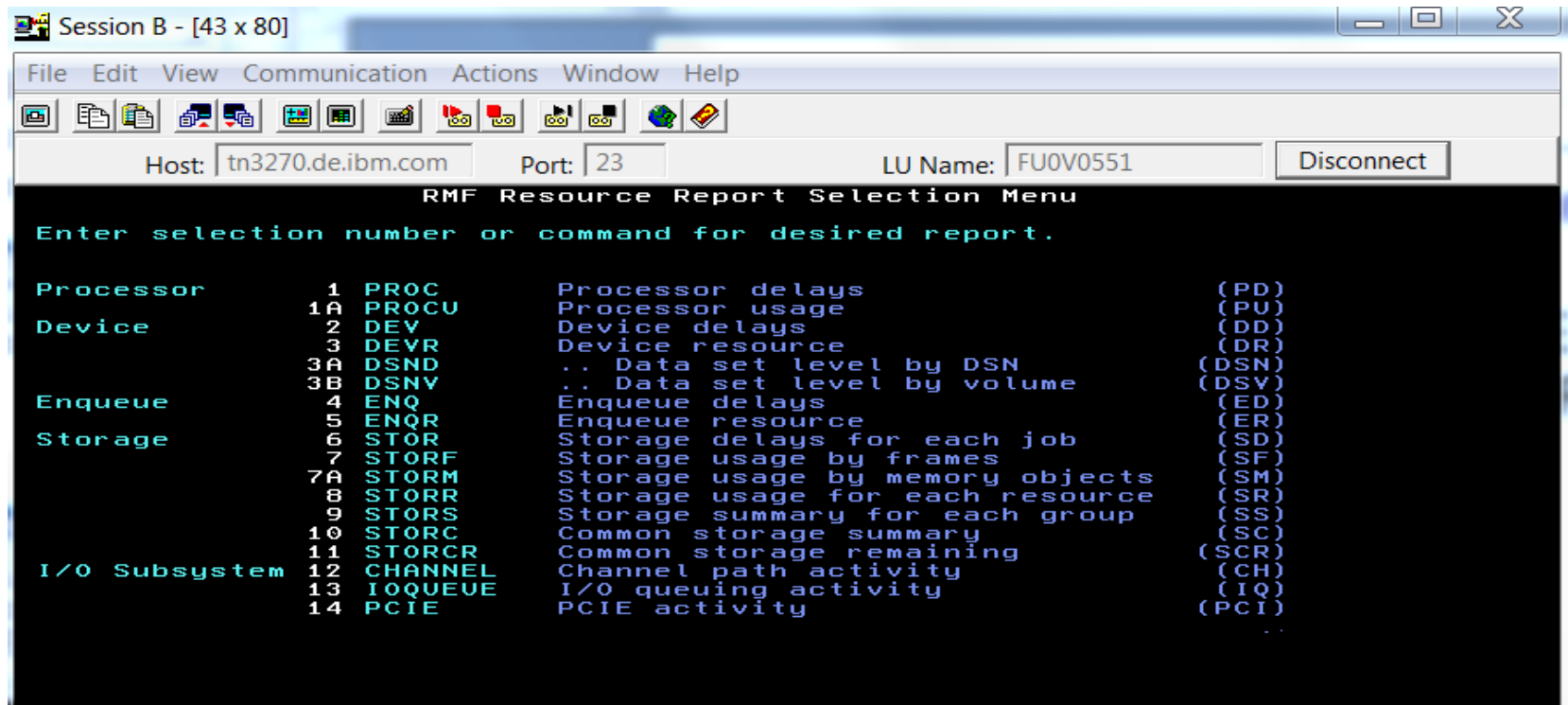
.....
VSAMRLS          /* ACTIVATE VSAM RLS GATHERING
OPD              /* ACTIVATE OMVS PROCESS DATA GATHERING
PCIE            /* ACTIVATE PCIE DATA GATHERING
NOZFS            /* NO ZFS DATA GATHERING
NOSGSPACE        /* NO STORAGE GROUP SPACE GATHERING
NOLOCK           /* NO LOCK DATA GATHERING

```



## Usage & Invocation: RMF Monitor III PCIE Activity Report


- To request the Monitor III PCIE Activity Report, select a 3 from the Primary Menu, then select 14 from the Resource Report Selection Menu or enter command: *PCIE* or *PCI*





## Usage & Invocation: RMF Monitor III PCIE Activity Report

- 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

■  Session B - [43 x 80]

File Edit View Communication Actions Window Help

Host: tn3270.de.ibm.com Port: 23 LU Name: FU0V0551 Disconnect

RMF V2R2 PCIE Activity Line 1 of 3

Samples: 60 System: NP0 Date: 08/21/14 Time: 19.54.00 Range: 60 Sec

ID	PCID	Type	Function Jobname	ASID	Status	Alloc Time%	PCI Load	Operations Store Block	Rate Refr	DMA Read	Rate Write
0023	05FC	zEDC	FPGHWAM	0013	Alloc	99.7	0	171	0 20.0	3909	3395
0024	0540	zEDC	FPGHWAM	0013	Alloc	99.7	0	171	0 19.9	215	287
0080	057C	RoCE	VTAM390	0029	Alloc	99.7	0.803	6140	0 0.669	2801	2011



## Usage & Invocation: PCIE Activity Report Main Panel Fields

Field Heading	Meaning
Function ID	Hexadecimal identifier of the PCIE Function for which performance data is reported.
Function PCID	Physical channel identifier for the PCIE function.
Function Type	Device type for the PCIE function which can be one of the following: <b>HWA</b> Hardware Accelerator <b>Oth</b> Unknown device type <b>RoCE</b> Remote Direct Memory Management <b>zEDC</b> zEnterprise Data Compression
Function Jobname	Name of the job who allocated the PCIE function.
Function ASID	Address space ID of the job who allocated the PCIE function.
Function Status	The PCIE function status at the end of this reporting interval: <b>Alloc</b> The function is allocated and in use <b>Dpend</b> The function is in the process of de-allocation <b>Error</b> The function is in permanent error <b>DeAlloc</b> The function is de-allocated <b>Unknown</b> The function status is unknown
Alloc Time %	Percentage of reporting interval for which the PCIE function was allocated or in the process of de-allocation.
PCI Operations Rate Load	Rate of PCI Load operations that were executed within this reporting interval.
PCI Operations Rate Store	Rate of PCI Store operations that were executed within this reporting interval.
PCI Operations Rate Block	Rate of PCI Store Block operations that were executed within this reporting interval.
PCI Operations Rate Refresh	Rate of Refresh PCI Translations operations that were executed within this reporting interval.
DMA Rate Read	Number of megabytes per second that were transferred by DMA reads from all defined DMA address spaces to the PCIE function.
DMA Rate Write	Number of megabytes per second that were transferred by DMA writes from the PCIE function to all defined DMA address spaces.



## Usage & Invocation: HW Accelerator And Compression Activity

- 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

Session B - [43 x 80]

File Edit View Communication Actions Window Help

Host: tn3270.de.ibm.com Port: 23

RMF V2R2 PC

Samples: 60 System: LPO Date: 08/21/14

ID	PCID	Type	Jobname	ASID	Status
0023	05FC	zEDC	FPGHWAM	0013	Alloc
0024	0540	zEDC	FPGHWAM	0013	Alloc
0080	057C	RoCE	VTAM390	0029	Alloc

RMF Hardware Accelerator And Compression Activity

Press Enter to return to the Report panel.

Function ID : 0024 Alloc Time % : 99.7 More: +  
Allocated : 19.55.00 on 08/21/14

Hardware Accelerator  
Time Busy % : 0.472 Transfer Rate : 8.02

Request  
Execution Time : 27.6 Std. Deviation: 7.50  
Queue Time : 14.8 Std. Deviation: 42.2  
Size : 47.0

Buffer Pool  
Memory Size : 16 Utilization : 0

	Compression	Decompression
Request Rate	170	0.585
Throughput	4.76	0.013
Ratio	1.45	0.419

F1=Help F2=SplitScr F3=End F6=RMFHelp  
F7=Backward F8=Forward F9=SwapScr F12=Return



## Usage & Invocation: HW Accelerator And Compression Activity Fields

Field Heading	Meaning
Allocated	Date and time when the Hardware Accelerator was allocated.
Hardware Accelerator Time Busy %	The percentage of time that this partition kept the Hardware Accelerator busy.
Hardware Accelerator Transfer Rate	The number of megabytes per second that were transferred by DMA operations.
Request Execution Time	The average time in microseconds the Hardware Accelerator took to process a request.
Request Execution Time Std. Deviation	The standard deviation of the request execution time.
Request Queue Time	The average queue time in microseconds that was spent for a request.
Request Queue Time Std. Deviation	The standard deviation of the request queue time.
Request Size	The average number of kilobytes that were transferred per request.
Buffer Pool Memory Size	The total size of memory in megabytes that is allocated to the buffer pool.
Buffer Pool Utilization %	The average utilization percentage of the buffer pool that z/OS kept for in-use buffers.
Request Rate	The number of compression or decompression requests per second.
Throughput	The number of megabytes that were compressed or decompressed per second.
Ratio	The ratio between input and output bytes that were compressed or decompressed within this reporting interval.



## Usage & Invocation: RoCE Device 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 device**

Session B - [43 x 80]

File Edit View Communication Actions Window Help

Host: tn3270.de.ibm.com Port: 23 LU Name: FU0V6106 Disconnect

RMF V2R2 PCIE Activity Line 1 of 3

Samples: 60 System: NP0 Date: 08/21/14 Time: 19.54.00 Range: 60 Sec

ID	PCID	Type	Function Jobname	ASID	Status	Alloc Time%	PCI Load	Operations Store	Block	Rate Refr	DMA Read	Rate Write
0023	05FC	zEDC	FPGHWAM	0013	Alloc	99.7	0	171	0	20.0	3909	3395
0024	0540	zEDC	FPGHWAM	0013	Alloc	99.7	0	171	0	19.9	215	287
0080	057C	RoCE	VTAM390	0029	Alloc	99.7	0.803	6140	0	0.669	2801	2011



## Usage & Invocation: DDS Full PCIE Activity Report Support

- 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=PCIE&resource=,sysname,MVS\\_IMAGE](http://hostname:8803/gpm/rmfm3.xml?report=PCIE&resource=,sysname,MVS_IMAGE)

File Edit View History Bookmarks Tools Help

RMF Data Portal RMF Monitor III Report

boesysf.boeblingen.de.ibm.com:8803/gpm/rmfm3.xml?report=PCIE&resource=%22SYSF.MVS\_IMAGE%22

Google

Most Visited IBM IBM

20141015144700

RMF Report [SYSF,MVS\_IMAGE] : PCIE (PCIE Activity)

Time Range: 10/15/2014 14:47:00 - 10/15/2014 14:48:00





Function Id	Function PID	Function Type	Jobname	ASID	Function Status	% Alloc Time	Allocation Date	Allocation Time	Load Operational Rate	Store Operational Rate	Store Block Operational Rate	Refresh Operational Rate	DMA Read Rate	DMA Write Rate	HWA Time % Busy	HWA Transfer Rate	Request Execution Time	Request Queue Time	Request Size	Request Execution Time StdDev	Request Queue Time StdDev	Compression Request Rate	Decompression Request Rate	Compression Throughput	Decompression Throughput	Compression Ratio	Decompression Ratio	Buffer Pool Memory Size	Buffer Pool % Utilization	# DMA AS	HWA Type
0022	037C	zEDC	FPGHW0012	Alloc	100	10/06/14	15.30.4	0	0.017	0	0.533	0	0	0.001	0.020	45.6	718	74.0	0.629	2.71	0.266	0	0.017	0	5.55		16	0	1	zCompr	
0026	037C	zEDC	FPGHW0012	Alloc	100	10/06/14	15.30.4	0	0	0.533	0	0	0	0	0						0	0	0	0			16		1	zCompr	
0029	03BC	zEDC	FPGHW0012	Alloc	100	10/06/14	15.30.4	0	0	0.533	0	0	0	0	0						0	0	0	0			16		1	zCompr	
002C	03BC	zEDC	FPGHW0012	Alloc	100	10/06/14	15.30.4	0	0.017	0	0.533	0	0	0.001	0.017	42.3	669	62.9	4.61	5.08	0.266	0	0.015	0	7.07		16	0	1	zCompr	



## Usage & Invocation: DDS Resource Model Enhancements

- 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

### Children of: SYSF,\*,I/O\_SUBSYSTEM

Icon	Resource	Metrics	Attributes	Res-Type
	<a href="#">SYSF,*,ALL_SSIDS</a>	<a href="#">Metrics</a>	N/A	ALL_SSIDS
	<a href="#">SYSF,*,ALL_LCUS</a>	<a href="#">Metrics</a>	N/A	ALL_LCUS
	<a href="#">SYSF,*,ALL_CHANNELS</a>	<a href="#">Metrics</a>	N/A	ALL_CHANNELS
	<a href="#">SYSF,*,ALL_VOLUMES</a>	<a href="#">Metrics</a>	N/A	ALL_VOLUMES
	<a href="#">SYSF,*,ZFS</a>	<a href="#">Metrics</a>	N/A	ZFS
	<a href="#">SYSF,*,PCIE</a>	<a href="#">Metrics</a>	N/A	PCIE

- Child resource PCIE\_FUNCTION is added to resource PCIE

### Children of: SYSF,\*,PCIE

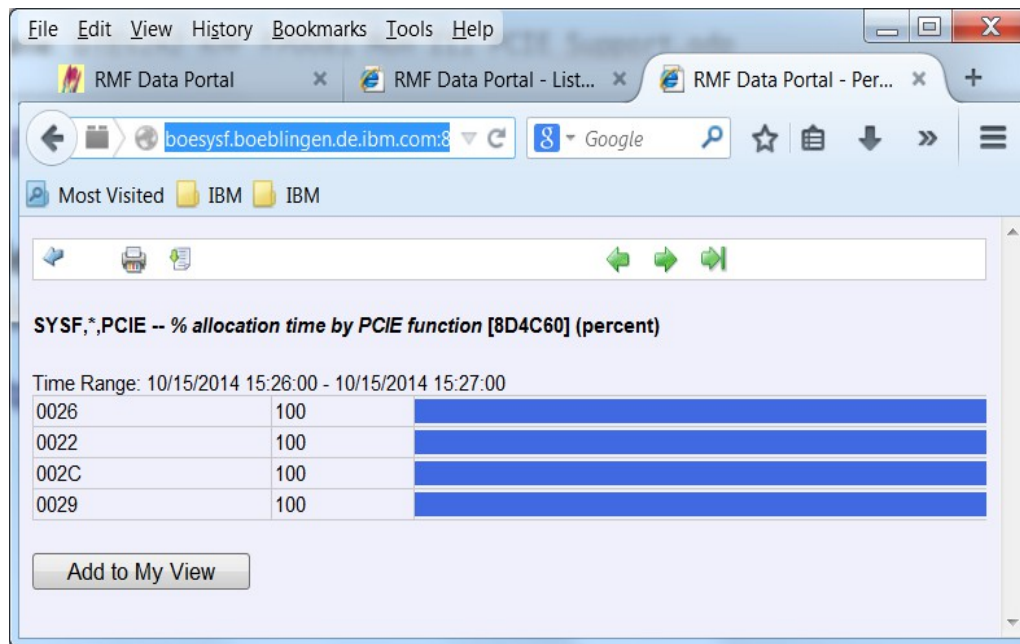
Icon	Resource	Metrics	Attributes	Res-Type
	<a href="#">SYSF,002C,PCIE_FUNCTION</a>	<a href="#">Metrics</a>	N/A	PCIE_FUNCTION
	<a href="#">SYSF,0022,PCIE_FUNCTION</a>	<a href="#">Metrics</a>	N/A	PCIE_FUNCTION
	<a href="#">SYSF,0026,PCIE_FUNCTION</a>	<a href="#">Metrics</a>	N/A	PCIE_FUNCTION
	<a href="#">SYSF,0029,PCIE_FUNCTION</a>	<a href="#">Metrics</a>	N/A	PCIE_FUNCTION



## Usage & Invocation: DDS PCIE metrics

- 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)
- Example: % Allocation Time by PCIE Functions allocated in z/OS system *sysname*

[http://hostname:8803/gpm/perform.xml?/resource=sysname,\\*,PCIE,&id=8D4C60](http://hostname:8803/gpm/perform.xml?/resource=sysname,*,PCIE,&id=8D4C60)



% allocation time by PCIE function  
 % buffer pool utilization by PCIE function  
 % time busy by PCIE function  
 buffer pool memory size by PCIE function  
 compression ratio by PCIE function  
 compression request rate by PCIE function  
 compression throughput by PCIE function  
 decompression ratio by PCIE function  
 decompression request rate by PCIE function  
 decompression throughput by PCIE function  
 request execution time by PCIE function  
 request execution time standard deviation by PCIE function  
 request queue time by PCIE function  
 request queue time standard deviation by PCIE function  
 request size by PCIE function  
 transfer rate by PCIE function  
 DMA read rate by PCIE function  
 DMA write rate by PCIE function  
 PCI load operations rate by PCIE function  
 PCI refresh operations rate by PCIE function  
 PCI store block operations rate by PCIE function  
 PCI store operations rate by PCIE function





## Presentation Summary

- Online monitoring of hardware features RoCE and IBM zEnterprise Data Compression (zEDC) was added to RMF in the V2R2 release
  - New Monitor III PCIE Activity report can be invoked in ISPF and through DDS API
  - DDS API exploiters can obtain PCIE related metric values
  - PCIE activity data is collected by RMF Monitor III if data gathering option PCIE is set
    -
  -



## Appendix

- RMF homepage: [www.ibm.com/systems/z/os/zos/features/rmf/](http://www.ibm.com/systems/z/os/zos/features/rmf/)
  - Product information, newsletters, presentations, etc.
  - Downloads
    - RMF Spreadsheet Reporter
    - RMF PM Java Edition
    - RMF Postprocessor XML Toolkit
- RMF email address: [rmf@de.ibm.com](mailto:rmf@de.ibm.com)
- Documentation and news
  - *RMF Report Analysis*, SC34-2665
  - *RMF User's Guide*, SC34-2664
  - Latest version of PDF files can be downloaded from:
    - <http://www.ibm.com/systems/z/os/zos/bkserv/>

