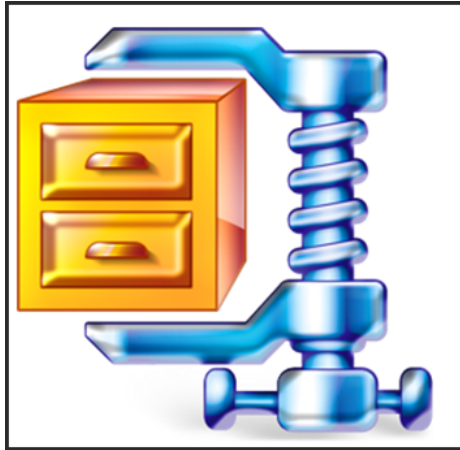


Glenn Anderson, IBM Lab Services and Training

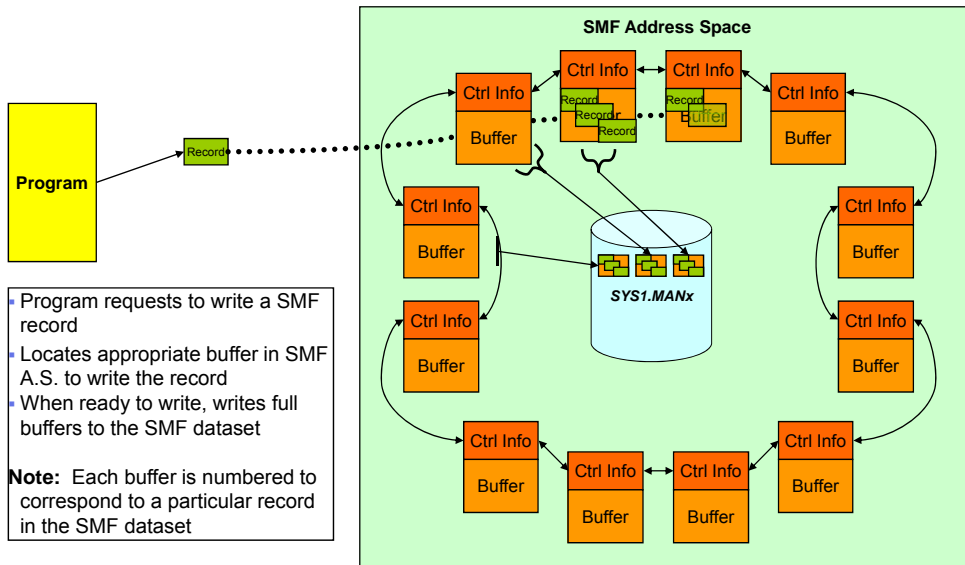


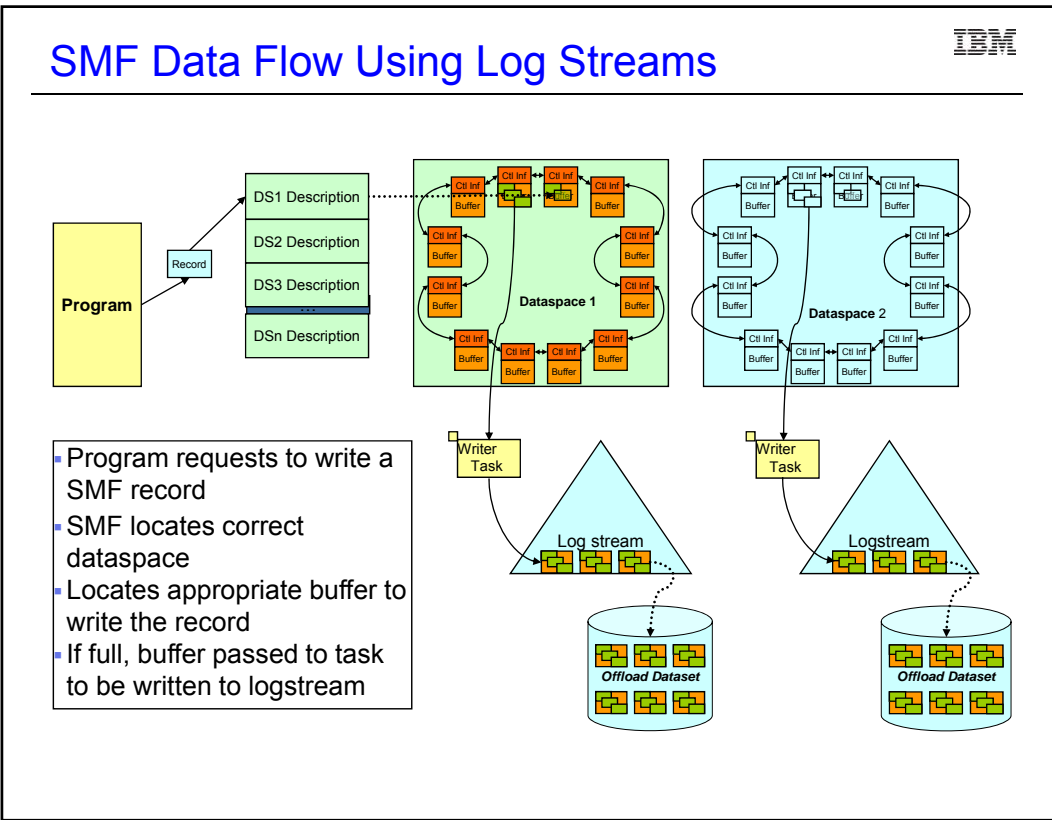
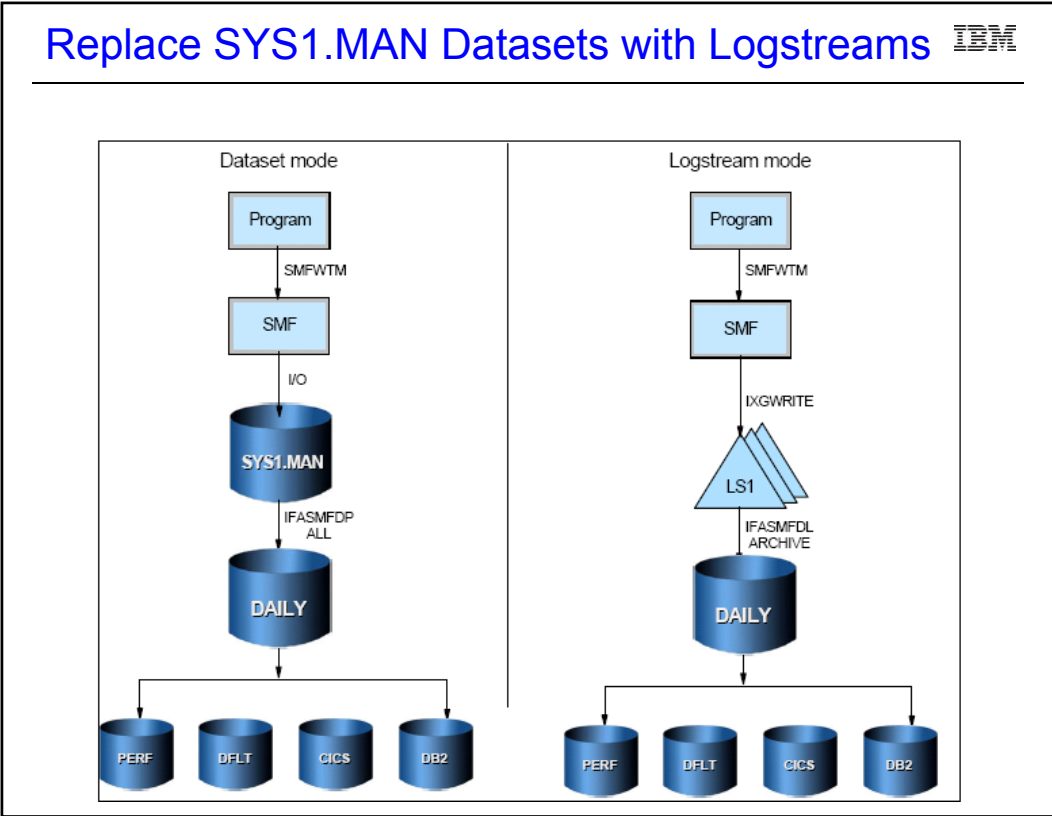
Configuring and Using SMF Logstreams with zEDC Compression

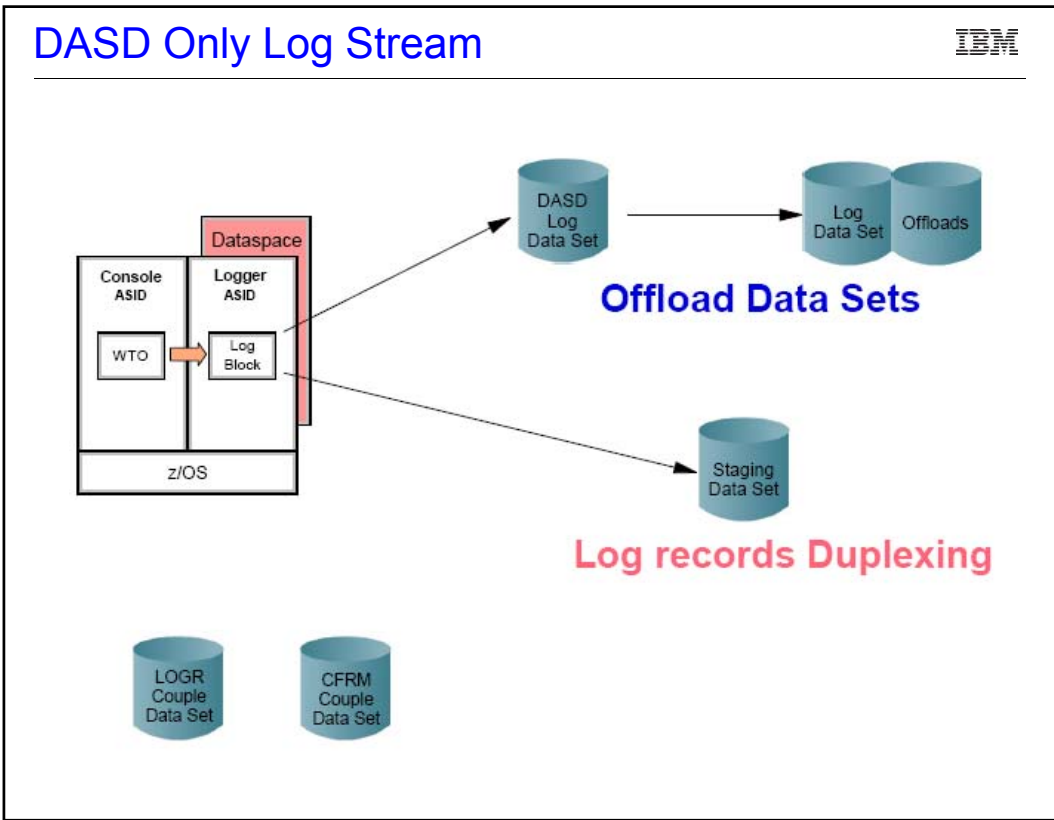
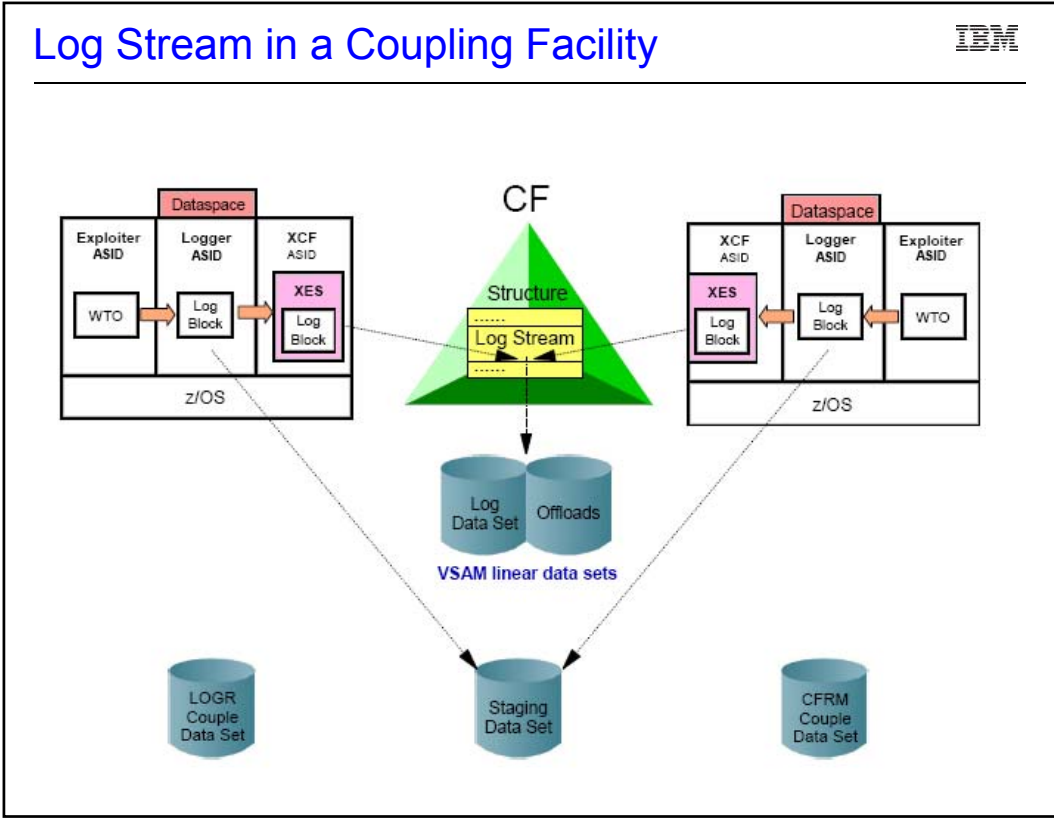


Summer SHARE
August 2015
Session 17644

Overview: Current SMF Data Flow









Usage and Invocation

- Define new logstreams in system logger
- Defining new keywords in SMFPRMxx
 - ▶ LSNAME(IFASMF.q1.q2,TYPE(xx:yy)
 - ▶ DEFAULTLSNAME(IFASMF.q1.q2)
 - ▶ RECORDING(DATASET|LOGSTREAM)
 - SETSMF operator command can be used to toggle recording settings
- Creating new JCL to use IFASMF DL with new logstreams
- Update processes to use data from logstreams, if necessary
- Activate PARMLIB changes via IPL or SET SMF=xx command



SMF Processing

- Relative data processing in IFASMF DL intended to mirror typical GDG processing
- **RELATIVEDATE** keyword
 - ▶ Specify DAILY, WEEKLY, or MONTHLY range and number of units
- IFASMF DL **LSNAME OPTIONS** to dump and/or delete data from logstream (vs. waiting for retention period to expire)
 - ▶ DUMP
 - ▶ DELETE
 - ▶ ARCHIVE (DUMP and DELETE)
- SMFPRMxx **MAXDORM** applies to SMF log streams (in addition to dataset recording)



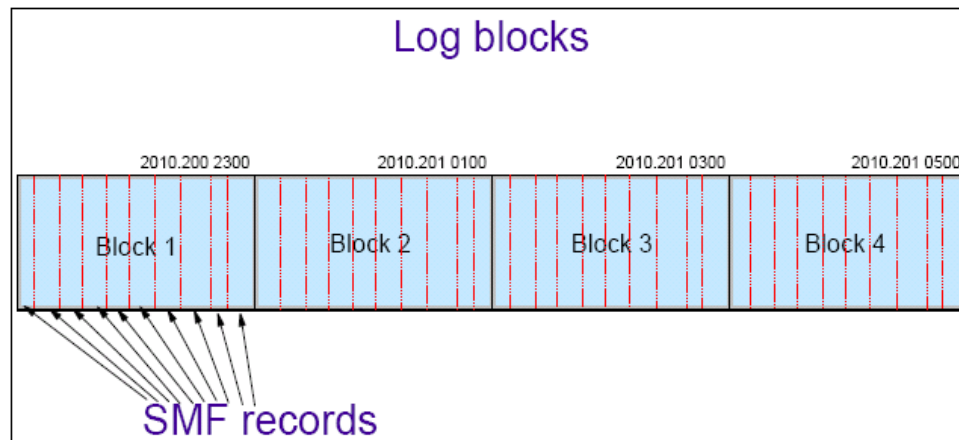
Usage and Invocation

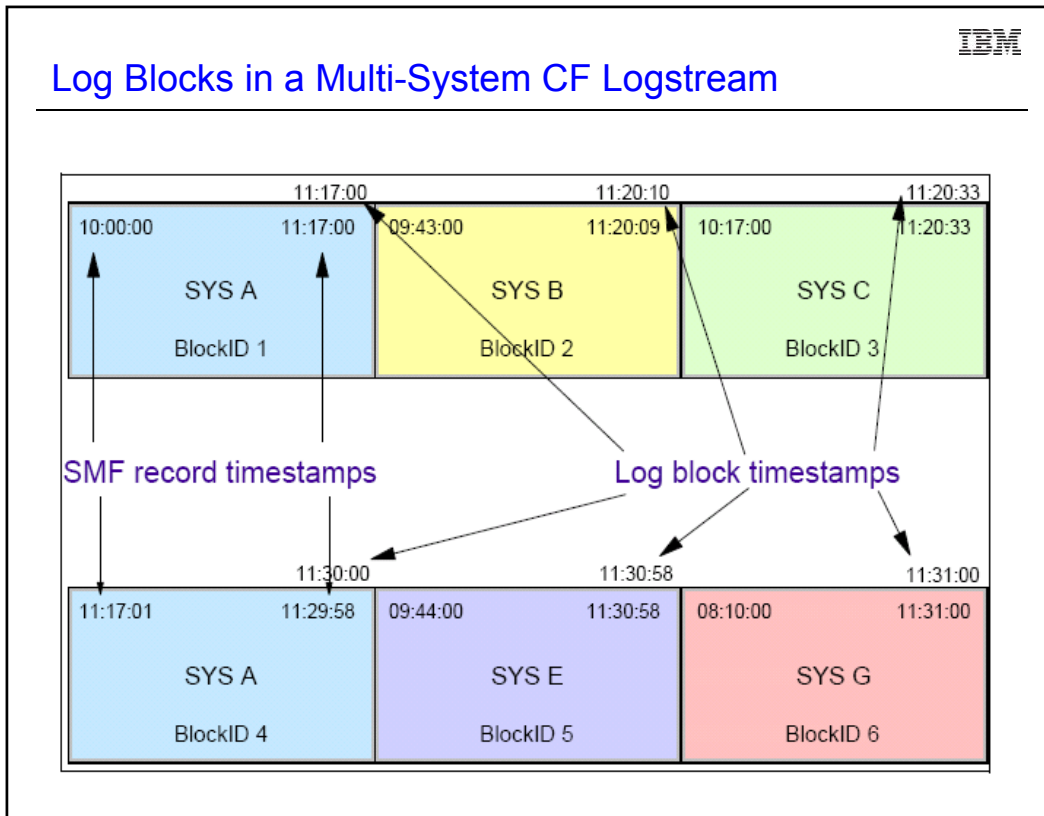
- The support for ARCHIVE, DELETE and RELATIVEDATE is invoked by the IFASMF DL program. The support for MAXDORM is invoked by updating your SMFPRMxx.
- RELATIVEDATE Parameter
 - ▶ Used to specify a date range based on the current day, week or month
 - **RELATIVEDATE(u, x, y)**
 - u – BYDAY, BYWEEK or BYMONTH
 - x – Number of units to move back
 - y – Number of units to gather
- DELETE/ARCHIVE Option
 - ▶ LSNAME(IFASMF.LS1,OPTIONS(ARCHIVE))
 - ▶ LSNAME(IFASMF.LS1,OPTIONS(DELETE))

9



Relationship of SMF Records to Log Blocks





IBM

IFASMFDDL Improvements in z/OS R13

- **Avoid reading to end of logstream**
 - ▶ IFASMFDDL starts reading a logstream at a point (approximately) representing a specified time
 - **SMARTENDPOINT** keyword to specify that IFASMFDDL should stop reading a logstream before the end
 - **SMARTEPOVER** specifies amount of time added to end date/time (default is two hours)
 - ▶ Avoids reading to end of logstream
- **Allow entire logstream to be archived or deleted**
 - ▶ Treat logstreams as though they were SMF datasets
 - ▶ Will reset logstream starting point to next new block



z/OS Ver 2.1 - SMF Logger Updates

- Specify log stream buffer sizes with new DSPSIZMAX parameter in SMFPRMxx
 - ▶ Support for DSPSIZMAX to be used when SMF is initialized also available for z/OS V1.12 and V1.13 with the PTF for APAR OA35175
 - ▶ z/OS V2.1 supports dynamic changes via SET SMF and SETSMF


- SMF also supports the use of data compression on zEC12 and zBC12 systems with the zEDC Express feature and the zEnterprise Data Compression (zEDC) feature for z/OS V2.1.





IBM z Enterprise Data Compression

What is it?	What Changes?	What is the Value?
<ul style="list-style-type: none"> ✓ <i>zEDC Express is an IO adapter that does high performance industry standard compression</i> ✓ <i>Used by z/OS Operating System components, IBM Middleware and ISV products</i> ✓ <i>Applications can use zEDC via industry standard APIs (zlib and Java)</i> ✓ <i>Each zEDC Express sharable across 15 LPARs, up to 8 devices per CEC.</i> ✓ <i>Raw throughput up to 1 GB/s per zEDC Express Hardware Adapter</i> 	<ul style="list-style-type: none"> ▪ Disk Savings: Many people are already getting value from CMPSC compression and software compression today ▪ Performance: High throughput alternative to existing System z compression for large or active files. ▪ Industry Standard: Low cost compressed data exchange across all platforms ▪ Pervasive: Standard APIs allow quick adoption by middleware products running on System z 	<ul style="list-style-type: none"> ▪ QSAM/BSAM compression can save disk cost ▪ Business Partner Data Exchange can have higher throughput with lower CPU cost ▪ Sterling Connect:Direct saves additional link bandwidth, elapsed time. ▪ ISV Products delivery expanded customer value ▪ Java transparently accelerates java.util.zip ▪ IBM Encryption Facility for standard compliant data exchange ▪ Improved availability with SMF compression

14

Compression Coprocessor (CMPSC) vs. zEDC 		
Compression Coprocessor	z Enterprise Data Compression	
<p>On Chip</p> <p>In every IBM eServer™ zSeries® today (and tomorrow)</p> <p>Mature: Decades of use by Access Methods and DB2®</p> <p>Work is performed jointly by CPU and Coprocessor</p> <p>Proprietary Compression Format</p>	<p>PCIe Adapter</p> <p>New with IBM zEnterprise® EC12 GA2 and IBM zEnterprise BC12</p> <p>Mature: Industry Standard with decades of software support</p> <p>Work is performed by the PCIe Adapter</p> <p>Standards Compliant (RFC1951)</p>	
<p>Use Cases</p>		
<p><u>Small object compression</u></p>	<p><u>Large Sequential Data</u></p>	<p><u>Industry Standard Data</u></p>
<ul style="list-style-type: none"> ▪ Rows in a database 	<ul style="list-style-type: none"> ▪ QSAM/BSAM Online Sequential Data ▪ Objects stored in a data base 	<ul style="list-style-type: none"> ▪ Cross Platform Data Exchange
<p><u>Users</u></p>	<p><u>Users</u></p>	<p><u>Users</u></p>
<ul style="list-style-type: none"> ▪ VSAM for better disk utilization ▪ DB2 for lower memory usage ▪ The majority of customers are currently compressing their DB2 rows 	<ul style="list-style-type: none"> ▪ QSAM/BSAM for better disk utilization and batch elapsed time improvements ▪ SMF for increased availability and online storage reduction ▪ DFSMSdss for better disk and tape utilization for backup data ▪ DFSMSHsm for improved CPU, disk reduction 	<ul style="list-style-type: none"> ▪ Java for high throughput standard compression via java.util.zip ▪ Encryption Facility for z/OS for better industry data exchange ▪ IBM Sterling Connect: Direct® for z/OS for better throughput and link utilization ▪ ISV support for increased client value
15		

zEDC Configuration Overview 
<ul style="list-style-type: none"> ▪ Operating system requirements <ul style="list-style-type: none"> • Requires z/OS 2.1 or later and new zEDC Express for z/OS feature <ul style="list-style-type: none"> –PTF Support for BSAM/QSAM • z/OS V1.13 and V1.12 offer software decompression support only ▪ Server requirements <ul style="list-style-type: none"> • Exclusive to z13, zEC12 and zBC12 • New zEDC Express feature for PCIe I/O drawer (FC#0420) <ul style="list-style-type: none"> •Each feature can be shared across up to 15 LPARs •Up to 8 features available on zEC12 or zBC12 and later • Recommended high availability configuration per server is four features <ul style="list-style-type: none"> •No additional software cost is incurred by increasing the number of hardware features •Provides high availability during concurrent update (half devices unavailable during update) •Recommended minimum configuration per server is two features • Hot pluggable • For best performance, feature is <i>needed on all systems accessing the compressed data</i> ▪ Capacity Planning <ul style="list-style-type: none"> • The z Batch Network Analyzer now reports on potential zEDC usage for QSAM/BSAM data sets
<p>16</p> <p>© 2015 IBM Corporation </p>

SMF Exploitation of zEDC Express



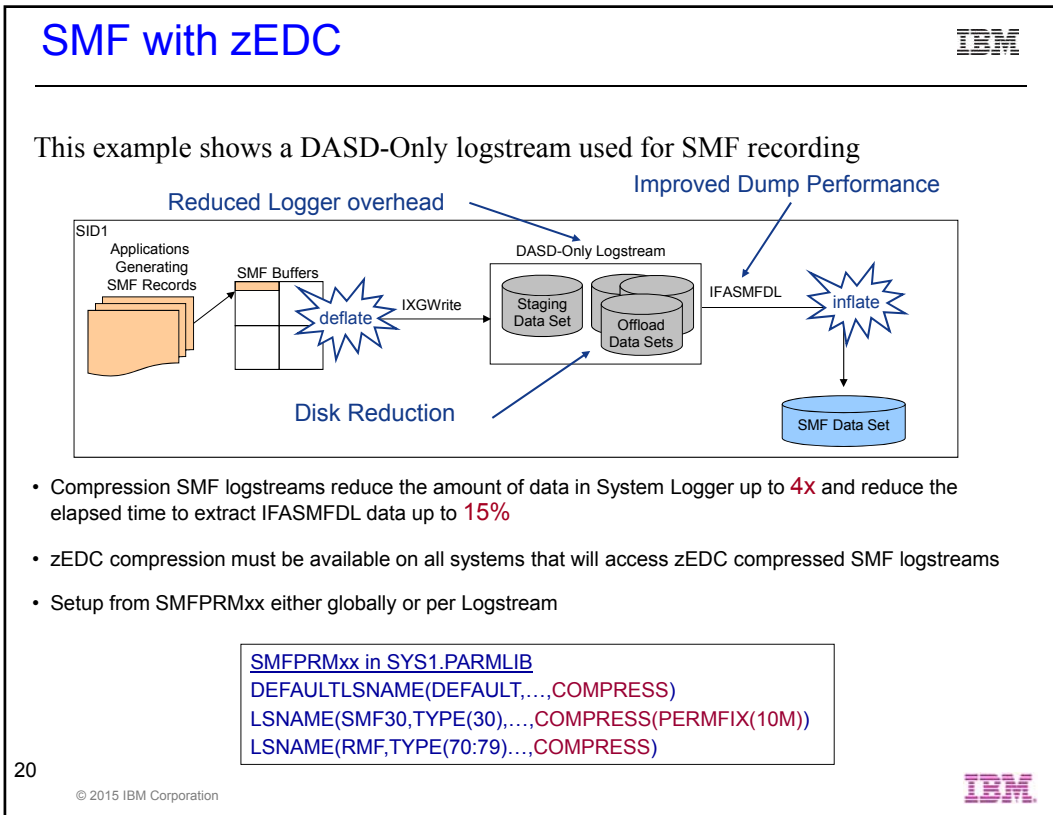
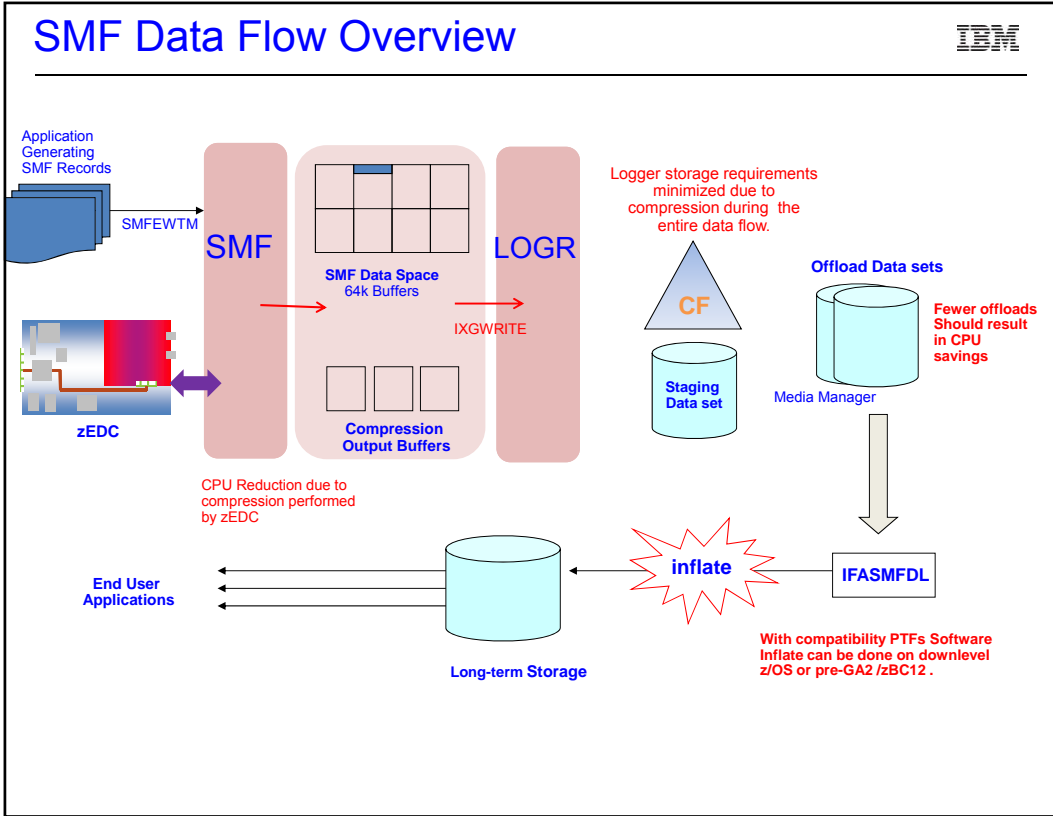
- SMF compression support is for both CF or DASD logstream
- Specify the new **COMPRESS** option on one or more logstream definitions
LSNAME or DEFAULTLSNAME
 - Option to specify amount of memory to permanently fix for performance
 - Same SMF records can be directed to multiple log streams and compression can be enabled on one of them
- IFASMF DL requirements
 - No changes required if zEDC is available; it will be used automatically
 - Specify the SOFTINFLATE option to process compressed data when there are no zEDC devices available
 - Requires z/OS PTF to provide software inflate (decompression) capability for z/OS 1.12 and 1.13 systems
 - If the SOFTINFLATE option is not specified on a system with no zEDC device, an error will occur and no records will be deleted from the SMF logstream
- Following SMF records to collect performance information:
 - SMF 23 – SMF buffer usage, number of records written etc.
 - SMF 88 - System logger log stream size, frequency of offload
 - Collection enablement is via SMFPRMxx

17

zEDC and SMF Logstream Data



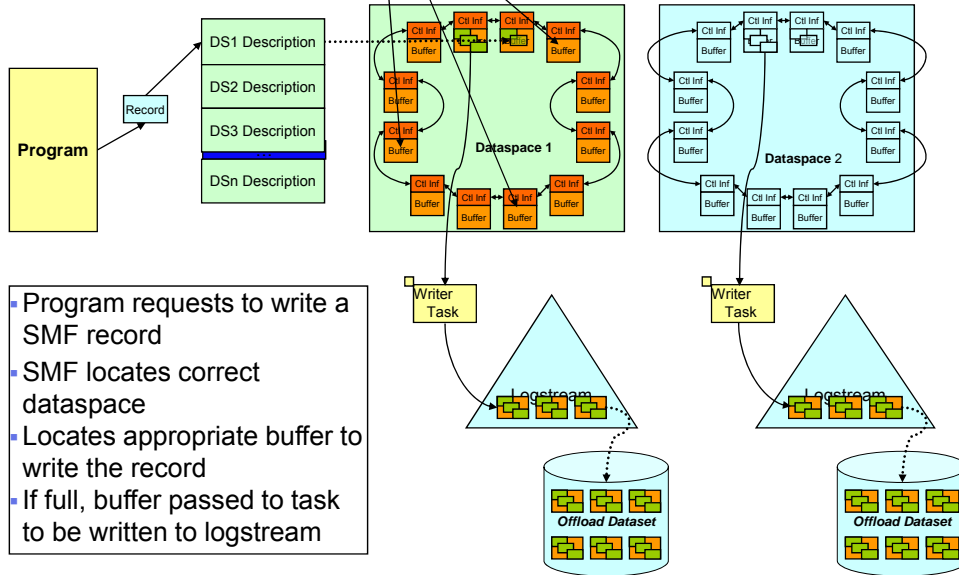
- New SMFPRMxx **COMPRESS** keyword on [LSNAME](#) and [DEFAULTLSNAME](#)
 - A buffer of SMF records is compressed by zEDC Express before it is written to the system logger
 - SMF data is only compressed while it is resident in the system logger
 - **PERMFI**X to specify amount of storage used for SMF buffers that can remain permanently fixed
- When compressed data is processed by IFASMF DL, it decompresses the SMF records for selection and writing
 - **SOFTINFLATE** parameter to process compressed SMF records using software algorithm, for a pre-z/OS V2.1 system or no zEDC Express



Logstream Buffer Params



DSPSIZMAX and PERMFIX



- Program requests to write a SMF record
- SMF locates correct dataspace
- Locates appropriate buffer to write the record
- If full, buffer passed to task to be written to logstream

Obtain PCIe Information by Command



- Use the console command **D PCIE** for general status information

```

D PCIE
I QP0221 12.06.02 DISPLAY PCIE 970
PCIE 0013 ACTIVE
PFID DEVICE TYPE NAME STATUS ASID JOBNAME PCHID VFN
0001 Hardware Accel erator ALLC 0014 FPGHWAM 0380 0001
0020 10GbE RoCE ALLC 00DE VTAM 038C
0021 10GbE RoCE CNFG 03B0
0011 Hardware Accel erator ALLC 0014 FPGHWAM 05C4 0001
    
```

Callouts in the diagram identify fields: **PCIe Started Task** (I QP0221), **PCIe Function ID** (0013), **Owner Name** (FPGHWAM), **Physical Channel Identifier** (0380), and **Virtual Function Number** (0001).

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

```

D PCIE,PFID=001
I QP0241 12.10.53 DISPLAY PCIE 521
PCIE 0013 ACTIVE
PFID DEVICE TYPE NAME STATUS ASID JOBNAME PCHID VFN
0001 Hardware Accel erator 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
    
```

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* and *IQPYPFMB PCIE Function Measurement Block*
- ▶ 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)

RMF Postprocessor PCIE Activity Report



RMF Postprocessor Interval Report [System Z2] : PCIE Activity Report

RMF Version : z/OS V2R1 SMF Data : z/OS V2R1
 Start : 00/13/2013-05:45:00 End : 00/13/2013-05:00:01 Interval : 15:00:000 minutes

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	Flash PCI Translations Operations Rate	DMA Address Space Count Rate	DMA Read Data Transfer Rate	DMA Write Data Transfer Rate
0001	0380	Hardware Accelerator: 1014044B	Allocated	FFQHWAM	0014	900	0	0.091	0	0	2.91	1	62.0	0
0011	05C4	Hardware Accelerator: 1014044B	Allocated	FFQHWAM	0014	900	0	0.091	0	0	2.52			
0020	038C	100BE RoCE	Allocated	VTAM	000E	900	0.889	0	0	0	0			

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
0001	0.005	31.4	4.88	545	08.0	75.2	0.110
0011	0.004	30.7	5.35	541	03.3	74.4	0.109

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
0001	1.46	0.088	4.11	0	0		64	0
0011	1.46	0.087	3.94	0	0		64	0

Callouts in the screenshot:

- Basic PCIe Metrics, e.g. PCI Load/Store and DMA Operations (points to the General PCIE Activity table)
- Common Request Statistics across all Personalities (Compression and future Personalities) (points to the Hardware Accelerator Activity table)
- Compression related Statistics (points to the Hardware Accelerator Compression Activity table)

▼ 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
0001	0380	Hardware Accelerator	1014044B	Allocated	FPGHWAM	0014	900	0	0.091
0011	05C4	Hardware Accelerator	1014044B	Allocated	FPGHWAM	0014	900	0	0.091
0020	038C	10GbE RoCE	15B31003	Allocated	VTAM	00DE	900	0.889	0

▼ Hardware Accelerator Activity

Function ID	Time Busy %	Request Execution Time	Std Dev for Request Execution Time	Request Queue Time	Std Dev for Request Queue
0001	0.005	31.4	4.88	545	68.0
0011	0.004	30.7	5.35	541	93.3

▼ Hardware Accelerator Compression Activity

Function ID	Compression Request Rate	Compression Throughput	Compression Ratio	Decompression Request Rate	Decompression Throughput
0001	1.46	0.088	4.11	0	0
0011	1.46	0.087	3.94	0	0

zEDC RMF Reporting IBM

New RMF report shows the utilization of each device.

RMF Postprocessor Interval Report : PCIE Activity Report

RMF Version : z/OS V2R1 SMF Data : z/OS V2R1
 Start : 02/24/2014-05.48.00 End : 02/24/2014-05.48.44 Interval : 00:45:000 minutes

▼ 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
0013	0.889	7.78	0.417	15.0	0.993	24.3	21.5

▼ Hardware Accelerator Compression Activity

Function ID	Compression Request Rate	Compression Throughput	Compression Ratio	Decompression Request Rate	Decompression Throughput	Decompression Ratio
0013	885	11.6	1.14	0	0	84

The percent of this interval where this specific zEDC Express device was executing requests

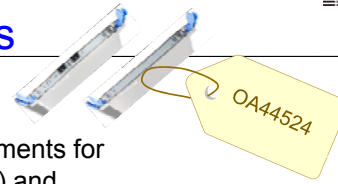
Compression ratio of all requests serviced by zEDC. This will span all users of this device.

Average request queue time in Microseconds for this device.

26 IBM

© 2015 IBM Corporation

z13 – RMF PCIE Enhancements



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

RMF Postprocessor PCIE Activity Report



RMF Postprocessor Interval Report [System Z2] : PCIE Activity Report

RMF Version : z/OS V2R1 SMF Data : z/OS V2R1
 Start : 08/13/2013-05:45:00 End : 08/13/2013-06:00:01 Interval : 15:00:000 minutes

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	Read Transfer Rate	Write Transfer Rate	Packets Received Rate	Packets Transmitted Rate	Work Units Processed Rate	Adapter Utilization
0001	0330	Hardware Accelerator	10140448	Allocated	FPQHWAM	0014	900	0	0.091	0	2.91	1	52.0	0		
0011	05C4	Hardware Accelerator	10140448	Allocated	FPQHWAM	0014	900	0	0.091	0	2.92	1	14.3	19.1		
0020	038C	10GbE RoCE	15B31003	Allocated	VTFAM	00DE	900	0.889	0	0	0	1	42.9	0		

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
0001	0.005	31.4	4.88	545	68.0	75.2	0.110	
0011	0.004	30.7	5.35	541	93.3	74.4	0.109	

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
0001	1.46	0.088	4.11	0	0		64	0
0011	1.46	0.087	3.94	0	0		64	0

RMF Postprocessor PCIE Activity Report

PCI Store Operations Rate	PCI Store Block Operations Rate	Refresh PCI Translations Operations Rate	DMA Address Space Count	Read Transfer Rate	Write Transfer Rate	Packets Received Rate	Packets Transmitted Rate	Work Units Processed Rate	Adapter Utilization
0.091	0	2.91	1	62.0	0				
0.091	0	2.92	1	14.3	19.1				
0	0	0	1	42.9	0				

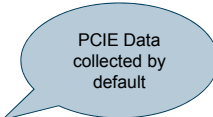
z/OS V2.2 – RMF Monitor III PCIE Activity Report

- 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



```

/*****
/* 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
NOZFS                /* NO ZFS DATA GATHERING
    
```



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

RMF V2R2 PCIE Activity

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

Function ID	PCID	Type	Jobname	ASID	Status	Alloc Time%	PCI Load	Operations Store	Rate Block	-Xfer Rate Refr	Read	Write
00A2	013C	RoCE	VTAM390	002C	Alloc	100	0.113	5999	0	1.69	0.267	288
006C	0204	zEDC	FPGHWAM	0013	Alloc	99.7	0	102	0	14.0		
007C	025C	zEDC	FPGHWAM	0013	Alloc							

RMF Hardware Accelerator And Compression Activity

Press Enter to return to the Report panel.

Function ID : 006C Alloc Time % : 99.7
 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

Buffer Pool
 Memory Size : 16 Utilization : 0

Request Rate : 102 Compression Decompression
 Throughput : 2.91 0.437
 Ratio : 2.79 0.009
 0.652

© Copyright IBM

Popup with extended Statistics for zEDC Accelerators

SMF 30 support for zEDC



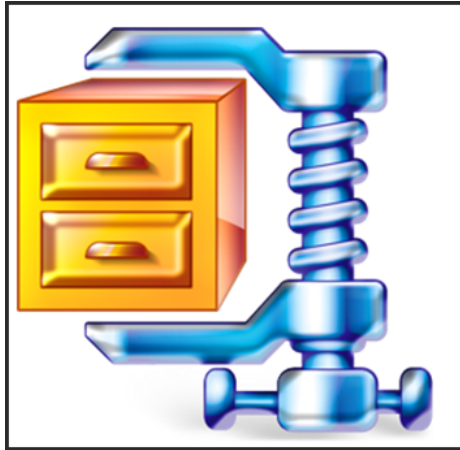
- Using zEDC, CPU time for job could go down
- How much zEDC capacity a job used?
- APAR OA45767 adds zEDC usage info to Type 30 records
 - Total number of zEDC requests
 - Total queue time in microseconds
 - Total execute time in microseconds
 - Deflate and inflate statistics for number of bytes input and output



Glenn Anderson, IBM Lab Services and Training



Configuring and Using SMF Logstreams with zEDC Compression



Thank you for attending!

Summer SHARE
August 2015
Session 17644