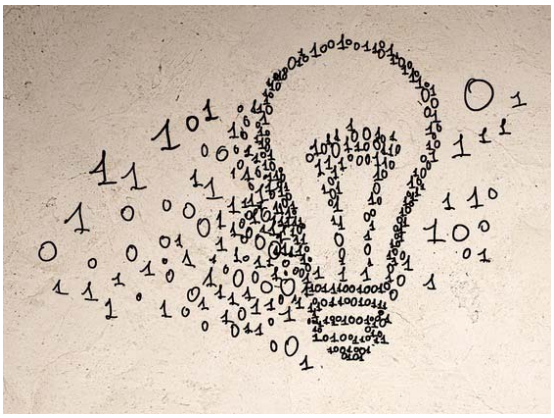
Glenn Anderson, IBM Lab Services and Training

Top New z/OS Performance Functions Every Sysprog Should Understand

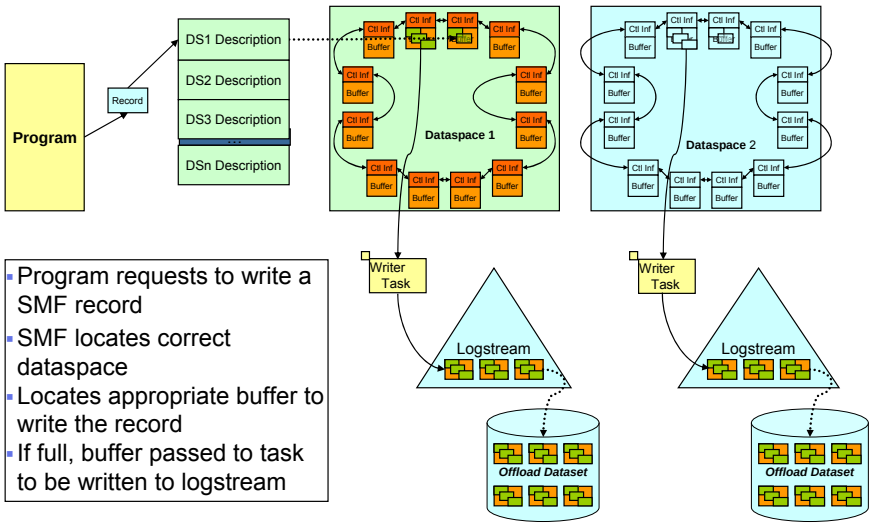


Winter SHARE
March 2014
Session 15219

Agenda

- SMF Logstreams and zEDC Support
- Flash Memory
- Warning Track
- Work-dependent Enclaves
- WLM and the IDAA
- RMF XP
- Some RMF Data I Like

SMF Data Flow Using Log Streams



- 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

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 IFASMFDL with new logstreams
- Update processes to use data from logstreams, if necessary
- Activate PARMLIB changes via IPL or SET SMF=xx command

SMF Processing

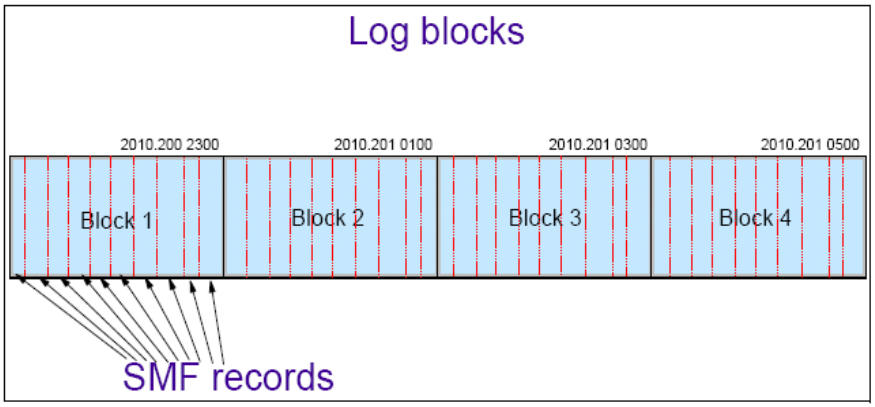
- Relative data processing in IFASMFDL intended to mirror typical GDG processing
- RELATIVEDATE keyword
  - Specify DAILY,WEEKLY,or MONTHLY range and number of units
- IFASMFDL LSNAME OPTIONS to dump and/or delete data from logstream (vs. waiting for retention period to expire)
  - DUMP
  - DELETE
 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

Usage and Invocation

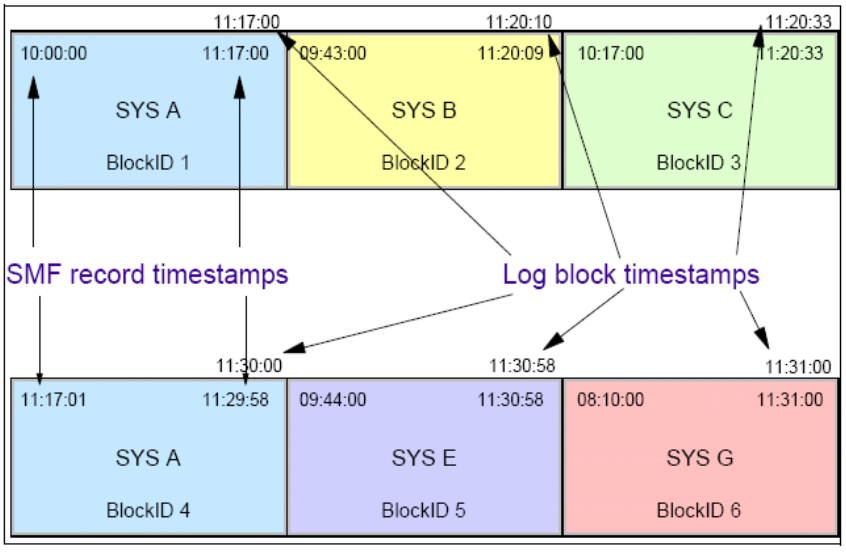
- The support for ARCHIVE, DELETE and RELATIVEDATE is invoked by the IFASMFDL program. The support for MAXDORM is invocated by updating your SMFPRMxx.
- RELATIVEDATE Parameter
  - Used to specify a date range based on the current day, week or month
    - **RELATIVEDATE(u, x, y)**
    - u – BYDAY, BYWEEK 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**))

6

Relationship of SMF Records to Log Blocks



Log Blocks in a Multi-System CF Logstream



IFASMFDL Improvements in z/OS R13

- Avoid reading to end of logstream
  - IFASMFDL starts reading a logstream at a point (approximately) representing a specified time
    - **SMARTENDPOINT** keyword to specify that IFASMFDL 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 zEnterprise Data Compression (zEDC)

What is it?

✓ *A combined software (z/OS V2.1) and hardware (zEDC Express) solution designed to help reduce resource consumption, disk utilization and optimize cross platform exchange of data*

How is it different

- **Performance:** Efficient alternative for larger files. Reduced CPU overhead for SMF jobs.
- **Efficient:** Optimized algorithms scan text to locate the re-use of phrases and refers back to earlier references
- **Industry Standard:** Compatible with open zlib based compression – widely used across all platforms
- **Economical:** Reduced DASD space requirements and improved effective bandwidth without significant CPU overhead***

15% reduction in elapsed time for SMF extraction with up to **40%** reduction for CPU time *

Logger overhead reduced by up to **30%** **

* When running an SMF extraction/dump against an SMF logstream with records compressed by zEDC
** The amount of data sent to an SMF logstream can be reduced by up to 75% using zEDC compression – reducing logger overhead
*** SOD for BSAM/QSAM access methods
All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Use cases for zEDC

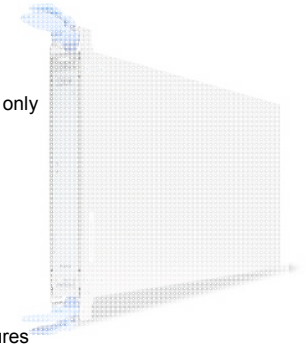
Use Cases

- **Reduced logger overhead** allows collection of more SMF data
- **Increase the amount of data you can keep active** by compressing more frequently accessed data
- **Enhance cross platform data exchange** when sending / receiving large data files
- **Improve disk utilization and economics of using flash** for extended format BSAM/QSAM
- **Improve latency for Java applications**

Target Market for zEDC

- Introductory Use with SMF Log Data: Clients running SMF using logger that are looking to reduce the logger overhead or collect additional data
- Clients such as a clearing house, financial institution or direct marketing agencies that are sending and receiving large files
- Customers with large volumes of extended format BSAM/QSAM sequential data
- Clients that have purchased flash on DS8870 and want to use it more efficiently when storing extended format BSAM/QSAM sequential data
- Clients that use Java today where they create a stream of compressed data

zEDC Requirements

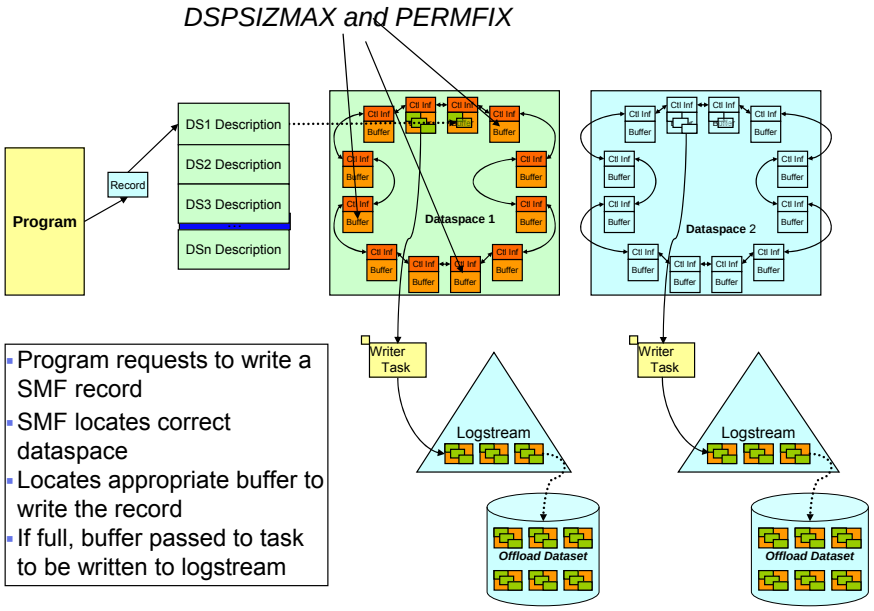


- **Operating system requirements**
  - Requires z/OS 2.1 and new zEDC for z/OS feature
  - z/OS V1.13 and V1.12 offer software decompression support only
  - Easy to set up and use – transparent to application software
    - Use policy (DATACLASS) to set up compression.
    - No changes to access method
- **Server requirements**
  - Exclusive to zEC12 (with Driver 15E) and zBC12
  - New zEDC Express feature for PCIe I/O drawer (FC#0420)
    - One compression coprocessor per zEDC Express feature
    - Each feature can be shared across up to 15 LPARs
  - Recommended minimum configuration per server is two features
    - Up to 8 features available on zEC12 or zBC12
  - For best performance, feature is ***needed on all systems accessing the compressed data***
- **Planned exploitation:**
  - Hardware exploitation first for log files - SMF records reduced logger overhead allows collection of more SMF data
  - All systems sharing sequential BSAM/QSAM extended format
  - Java using standard zlib compression library for compression services. Java applications and middleware can be transparently accelerated by enabling Java for hardware compression

zEDC and SMF Logstream Data

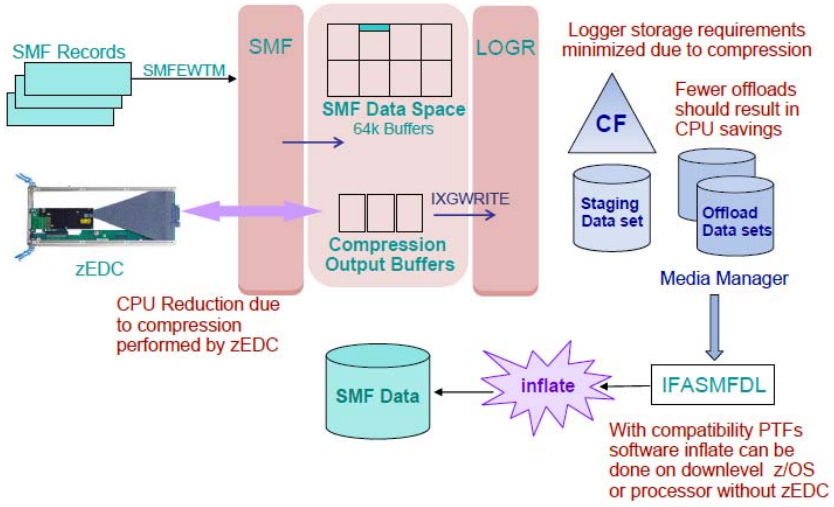
- New SMFPRMxx **COMPRESS** keyword on LSNAME and DEFAULTLSNAME
  - A buffer of SMF records is compressed by zEDC Express before it is written to the system logger
  - SMF data is only compressed while it is resident in the system logger
  - **PERMFIX** to specify amount of storage used for SMF buffers that can remain permanently fixed
- When compressed data is processed by IFASMFDL, 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 Parms



- 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

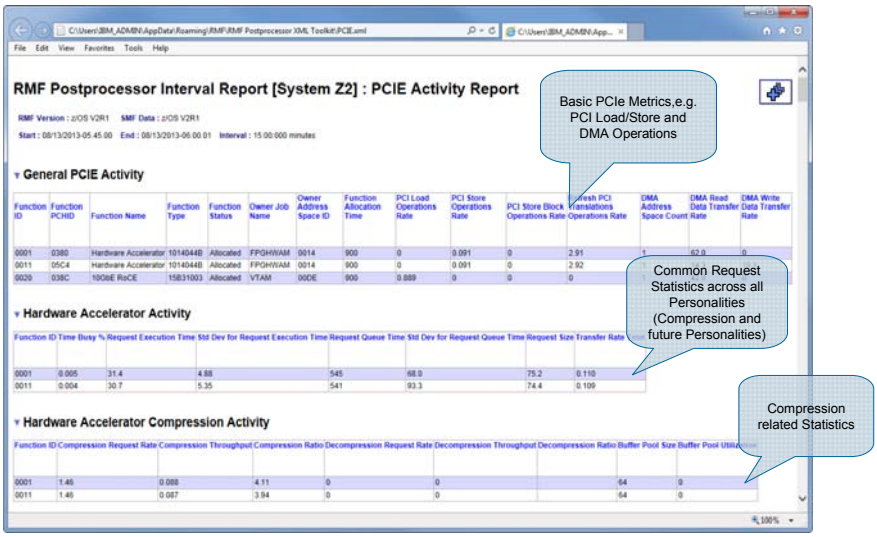
SMF Logstream Data Flow and zEDC



Obtain PCIe Information via API

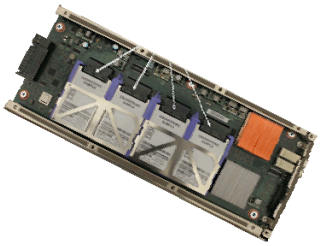
- IQPINFO – Obtain PCIe Information
  - The IQPINFO service provides PCIe related information, including any performance statistics
  - The service is described in *MVS Programming: Authorized Assembler Services Reference*
  - The response data area of the IQPINFO service is mapped by the macros
    IQPYPERF PCIE Performance Data Return Area
    IQPYPFMB PCIE Function Measurement Block
- 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

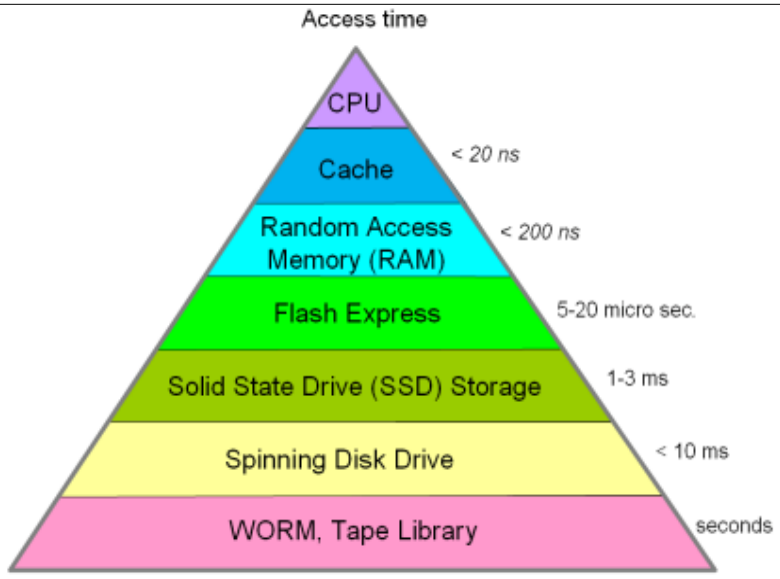


What is Flash Express?

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

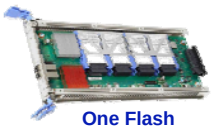


zEC12 – Flash Express



Flash Express Removes Last Vestiges of Unavailability

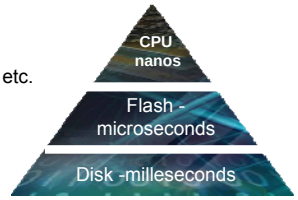
FLASH Express
- Unique application of Flash SSDs to server side
- Uses standard PCIe IO adapter. Physically comprised of internal SSDs on the card



One Flash Express Card

- **Capacity**
  - Sized large enough so that no capacity planning is needed
  - Can accommodate *all paging*
    - Each **card pair** provides **1.4 TB** storage; Maximum 4 card pairs (5.6 TB)
    - Typical customer has 6 - 8 LPARs per CEC and 40GB - 80GB for paging dataset size
  - Supported on z/OS V1.13 as well plus web deliverable

- **Qualities of Service**
  - Error Isolation, Transparent mirroring, Centralized diagnostics, etc.
  - Hardware Logging, FRU Call, Recovery
  - Concurrent Firmware update for service
  - **Immediately usable**
  - Minimal capacity planning needed
  - No intelligent data placement needed
  - Now dynamically reconfigurable
- **Secured**
  - Adapter is protected with 128-bit AES encryption.
  - Uses crypto hardware for secured data



Access Time

21

RSM Enhancements

- RSM Enhancements were delivered via RSM Enablement Offering Web Deliverable (FMID JBB778H) for z/OS V1.13
  - Exploit Storage Class Memory (SCM) technology for z/OS paging and SVC dump
    - Is expected to yield substantial improvements in SVC dump data capture time
    - Remove the requirement for non-VIO local page data sets when the configuration includes enough SCM to meet peak demands
    - However, local page data sets remain required for VIO, and when needed to support peak paging demands that require more capacity than provided by the amount of configured SCM
  - Pageable 1MB Large Page Support
  - Dynamic reconfiguration support for Storage Class Memory (SCM)
  - Optional PLPA and COMMON page data set support
  - 2GB Large Page Support

22

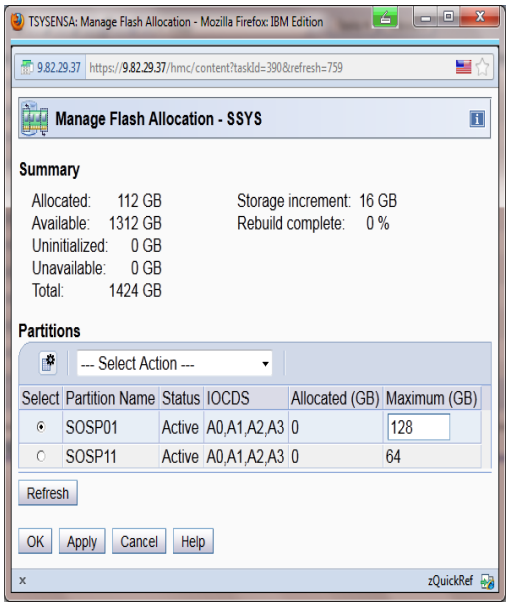
RSM Enhancement Considerations

- Installation of the z/OS V1R13 RSM Enablement Offering Web Deliverable (JBB778H) will:
  - Increase the size of the Nucleus by ~380K above the 16MB line
    - You may need to analyze your private area storage usage
  - Increase of 24K bytes (6 Pages) in ESQA per CPU per LPAR
    - This increase in ESQA per CPU includes CPs, zIIPs and zAAPs
  - New memory pool (Pageable Large Page) is automatically carved out (approximately 1/8 of above the bar real storage)
    - Converted to Pageable 4K Pages if needed by the system

Flash Allocation

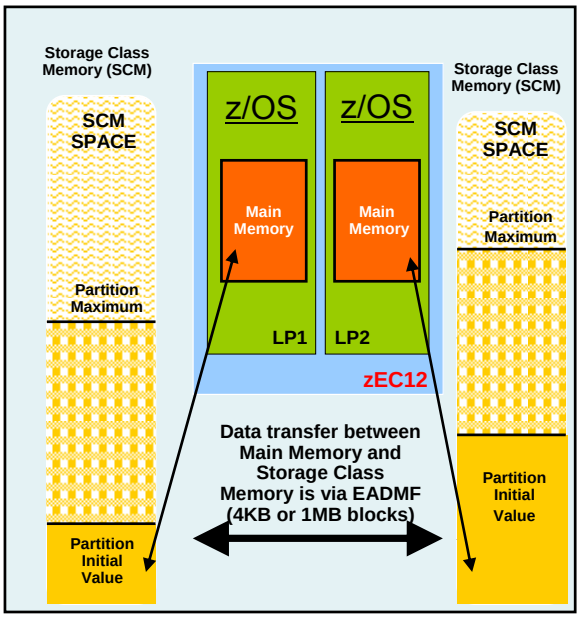
Allocating Flash to a partition

- The ***initial*** and ***maximum*** amount of Flash Memory available to a particular logical partition is specified at the SE or HMC via a new Flash Memory Allocation panel
- Can dynamically change maximum amount of Flash Memory available to a logical partition
- Additional Flash Memory (up to the maximum allowed) can be configured online to a logical partition dynamically at the SE or HMC
  - For z/OS this can also be done via an operator command
- Can dynamically configure Flash Memory offline to a logical partition at the SE or HMC
  - For z/OS this can also be done via an operator command
- Predefined subchannels, no IOCDS



Virtualization

- **Full virtualization of physical Flash PCIe cards across partitions, software sees an abstracted Flash Storage Space…**
  - Allows each logical partition to be configured with its own SCM storage space
  - Allocate Flash to partitions by amount, not card size
  - Ability to change underlying technology while preserving API



25

Representative Use Cases - Flash Express

Flash Express can reduce latency delays from paging to bring system availability to new heights and improve overall service levels

Application related errors will require collection of diagnostics. These diagnostics can be collected faster with Flash Express, reducing paging related delays that can impact your overall system availability.

Having your working data resident in Flash can help accelerate start of day processing, and improve service for many industries at the busiest time of their work day- a time when they cannot afford disruptions.

DB2 and Java in memory buffer pools work to store and process application data. DB2 and Java can benefit from 1MB pageable large pages with Flash Express, improving overall performance.

26

Flash for z/OS Paging Value

- Flash memory is a faster paging device as compared to HDD
  - The value is NOT in replacing memory with flash but replacing disk with Flash
  - Flash is suitable for workloads that can tolerate paging and will not benefit workloads that cannot afford to page
  - The z/OS design for flash memory does not completely remove the virtual storage constraints created by a paging spike in the system. (Some scalability relief is expected due to faster paging I/O with flash memory.)

27

z/OS Configuration and Setup

- New PAGESCM= keyword in IEASYSxx defines the amount of flash to be reserved for paging
  - Value may be specified in units of M, G, or T
  - NONE indicates do not used flash for paging
  - ALL is the default
    - Default indicates all flash defined to the partition is available for paging

28

z/OS V1.13 1 MB Pageable Large Page Exploitation

- Benefits of large pages:
  - Better performance by decreasing the number of TLB misses that an application incurs
  - Less time spent converting virtual addresses into physical addresses
  - Less real storage used to maintain DAT structures
- Fixed large pages vs pageable large pages:
  - Fixed large pages are backed at allocation. Pageable large pages are backed when referenced.
  - Use of fixed large pages for unauthorized users is controlled by a RACF profile (IARRSM.LRPAGES). No RACF authorization to use pageable large pages.
  - Fixed large pages stay as 1 MB pages while pageable large pages may be demoted to 4K pages in certain situations.
- Performance:
  - Java: performance with pageable 1MB large pages is equivalent to 1MB fixed large pages for java heap: up to 5% ITR impact
  - IMS using pageable large pages: up to 1% system ITR improvement. Expect more with z/OS V2.1.
  - DB2 using pageable large pages: up to 3% system ITR improvement.

CF Flash Initial Exploitation

- Initial CF Flash exploitation is targeted for MQ shared queues structures
  - Provides standby capacity to handle MQ shared queue buildups during abnormal situations, such as where “putters” are putting to the shared queue, but “getters” are transiently not getting from the shared queue
- Flash memory in the CEC is assigned to a CF partition via hardware definition panels, just like it is assigned to the z/OS partitions
- CFRM policy definition *permits* the desired maximum amount of Flash memory to be used by a particular structure, on a structure-by-structure basis
  - Note that Flash memory is NOT pre-assigned to structures at allocation time
- Structure size requirements for *real memory* get somewhat larger at initial allocation time to accommodate additional control objects needed to make use of Flash memory
- CFSIZER's structure recommendations will take these additional requirements into account, both for sizing the structure’s Flash usage itself, and for the related real memory considerations

30

zEC12 – Flash Memory & Pagegable Large Pages RMF Support

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

zEC12 – Flash Memory & Pagegable Large Pages

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

The number of shared pages backed on SCM

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

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

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

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

zEC12 – Flash Memory & Pagegable Large Pages...

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

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

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

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

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

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

HIGH COMMON FRAMES          TOTAL                  K   AUX DASD    AUX SCM
------------------  ---------                  -   ---------  ---------
      MIN               17301504                3           0          0
      MAX               17301504                3           0          0
      AVG               17301504                3           0          0
```

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

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

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

zEC12 – Flash Memory & Pagegable Large Pages...

- New SCM statistics RMF Postprocessor Page Data Set Activity report

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

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

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

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

zEC12 – Flash Memory & Pagegable Large Pages...

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

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

                                        PRIVATE AREA DETAIL

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

MEMORY ALLOCATION IN HIGH VIRTUAL MEMORY (ABOVE 2GB)

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

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

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

Decomposition into Fixed and Pageable Large Pages

zEC12 – Flash Memory & Pagegable Large Pages...

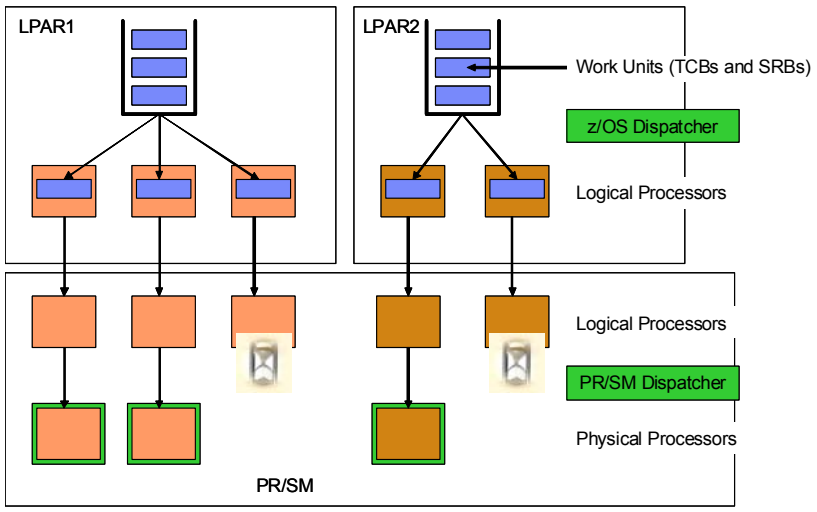
New Large Page statistics in RMF Monitor III STORM report

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

```
                          RMF V1R13  Storage Memory Objects

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

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

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

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

Dispatching in an LPAR environment: **Warning Track**

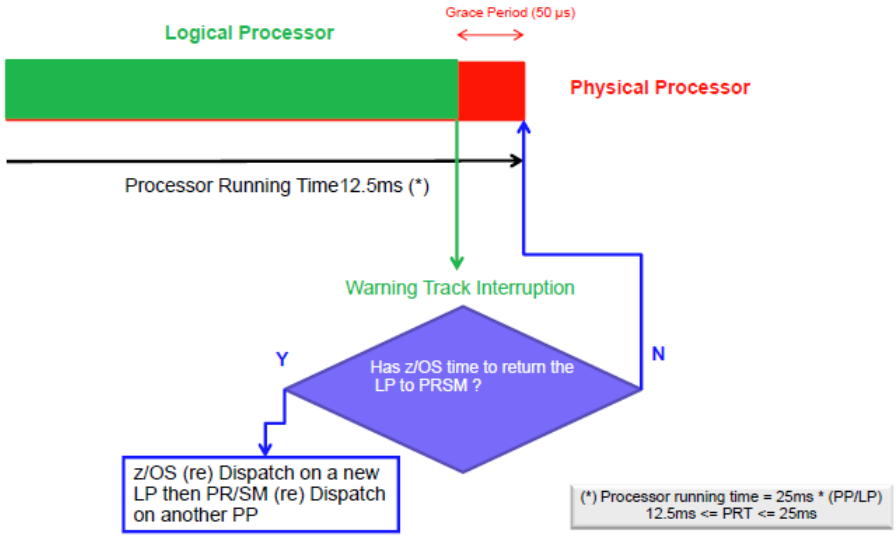


Warning track

IBM

- In a PR/SM™ environment the LPAR hypervisor assigns physical engines to logical engines accordingly to the weighting factors of the partitions.
- Once the time slice for a logical engine is expired the currently executing work is suspended until a physical engine is assigned to the logical engine again.
- The Warning Track Interruption Facility notifies the operating system that PR/SM™ will undispatch a certain logical processor within the next 50 microseconds (grace period).
- z/OS is now able to save status for the running unit of work and re-dispatch the work unit on a different logical processor within the grace period.
- z/OS now signals to PR/SM via Diagnose x'9C' that the logical processor can be un-dispatched.
- Warning Track processing is only supported in HyperDispatch=YES environments.
- A high benefit can be achieved for Low Share processors which might be parked by WLM.

Warning track



Latent Demand: LPAR Busy vs MVS Busy

CPU 2097 CPC CAPACITY 1451
MODEL 719 CHANGE REASON=N/A HIPERDISPATCH=YES

---CPU---		ONLINE	LPAR BUSY	MVS BUSY	PARKED	LOG PROC SHARE %	
NUM	TYPE		TIME %				
0	CP	100.00	96.77	96.80	0.00	100.0	HIGH
1	CP	100.00	94.91	94.95	0.00	100.0	HIGH
2	CP	100.00	96.72	96.74	0.00	100.0	HIGH
3	CP	100.00	95.07	95.10	0.00	100.0	HIGH
4	CP	100.00	50.18	93.55	0.00	66.0	MED
5	CP	100.00	50.15	93.56	0.00	66.0	MED
6	CP	100.00	20.30	89.09	56.00	0.0	LOW
7	CP	100.00	11.40	90.19	72.00	0.0	LOW
8	CP	100.00	22.12	88.49	50.79	0.0	LOW
9	CP	100.00	46.12	87.87	0.00	0.0	LOW
A	CP	100.00	45.37	86.74	0.00	0.0	LOW
B	CP	100.00	38.46	86.76	11.21	0.0	LOW
C	CP	100.00	35.08	86.96	19.43	0.0	LOW
D	CP	100.00	19.29	84.13	57.66	0.0	LOW
E	CP	100.00	0.00	-----	100.00	0.0	LOW
F	CP	100.00	0.00	-----	100.00	0.0	LOW
10	CP	100.00	0.00	-----	100.00	0.0	LOW
TOTAL/AVERAGE			42.47	91.45		532.0	

CEC Busy = 98.85

.0115 * 19 CP = .22 CPs available

Weight: 5.32 CPs

Using: 42.47/100 * 17 LCP = 7.22 CPs

Warning track statistics

IBM

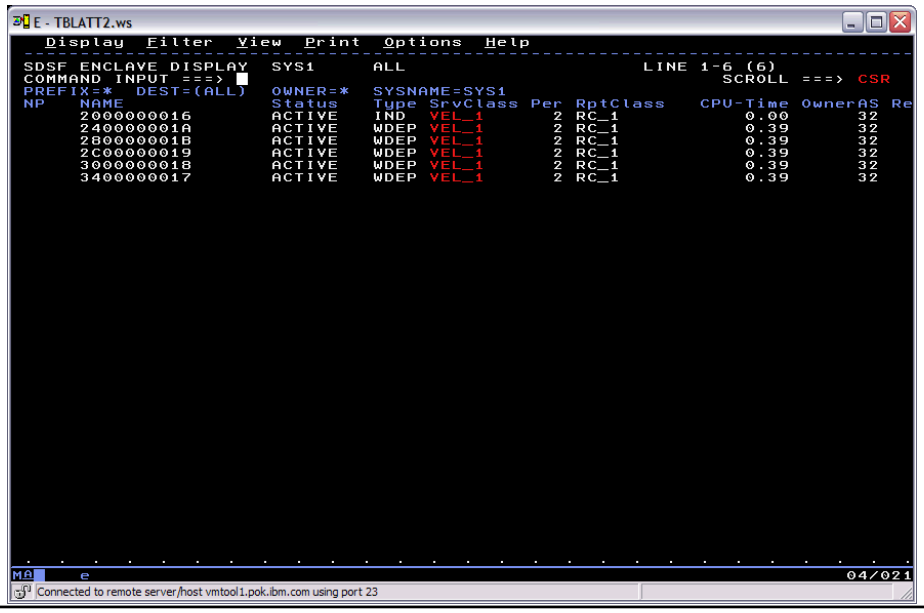
- RMF keeps track of the number of times PR/SM issued a warning-track interruption to a logical processor and z/OS was able/unable to return the logical processor within the grace period.
- RMF measures the amount of time in microseconds that a processor was yielded to PR/SM due to Warning-track processing.

SMF record type 70 subtype 1 (CPU Activity) – CPU data section				
Offset	Name	Length	Format	Description
80 x50	SMF70WTS	4	Binary	The number of times PR/SM issued a warning-track interruption to a logical processor and z/OS was able to return the logical processor within the grace period.
84 x54	SMF70WTU	4	Binary	The number of times PR/SM issued a warning-track interruption to a logical processor and z/OS was unable to return the logical processor within the grace period.
88 x58	SMF70WTI	4	Binary	Amount of time in microseconds that a logical processor was yielded to PR/SM due to Warning Track processing.

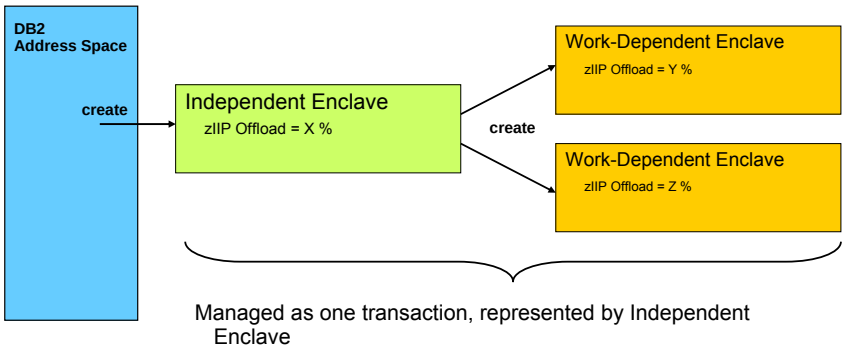
RMF Postprocessor Overview Conditions		
Name	Qualifier	Description
WTRKCP (WTRKAAP) (WTRKIIP)	cpu-id	The percentage of times PR/SM issued a warning-track interruption to a processor and z/OS was able to return it to PR/SM within the grace period.
WTRKTCP (WTRKTAAP) (WTRKTIIP)	cpu-id	Time in microseconds that a purpose processor was yielded to PR/SM due to Warning Track processing.

Work-dependent enclaves in SDSF

IBM

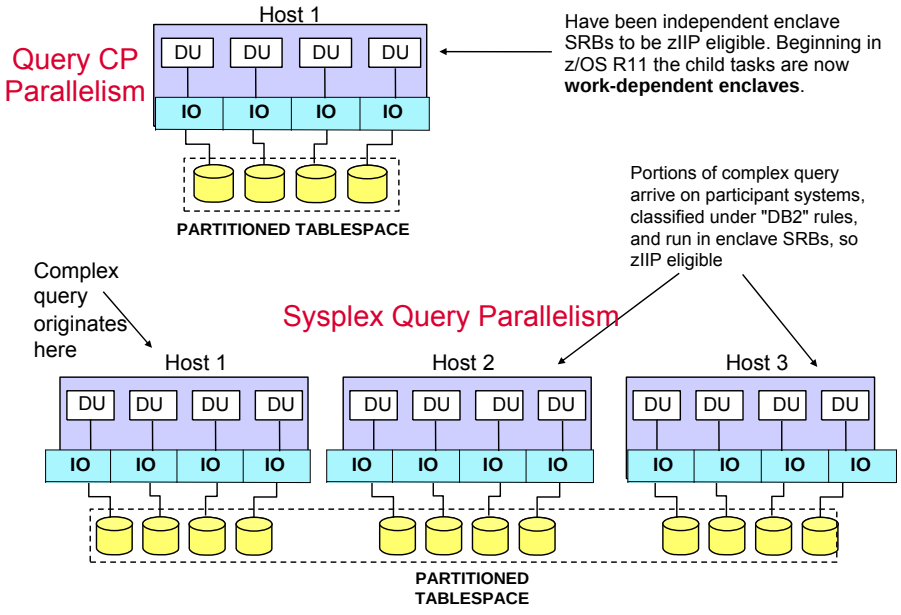


Work-Dependent enclaves



Implement a new type of enclave named “Work-Dependent” as an extension of an Independent Enclave. A Work-Dependent enclave becomes part of the Independent Enclave’s transaction but allows to have its own set of attributes (including zIIP offload percentage)

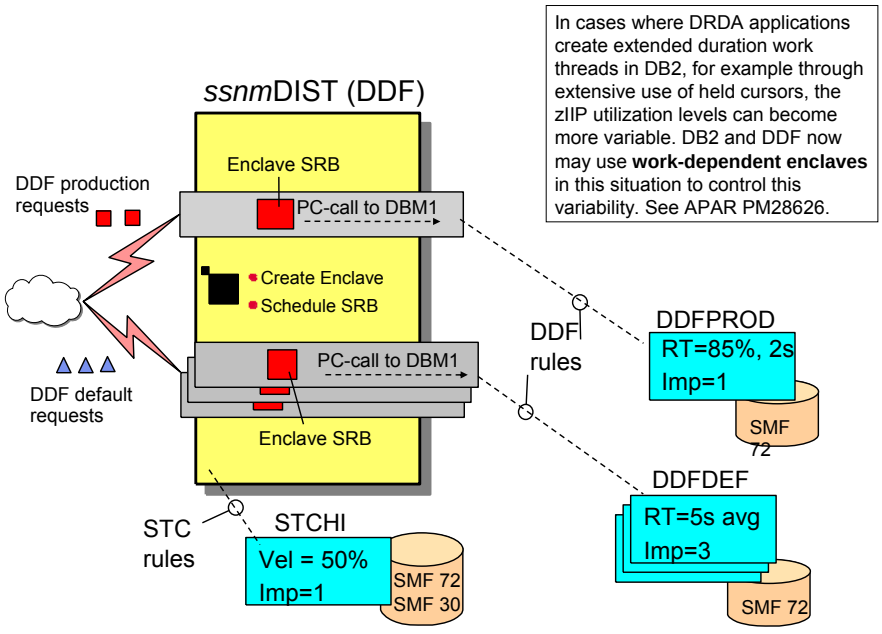
DB2 parallel query, enclave SRBs and zIIPs



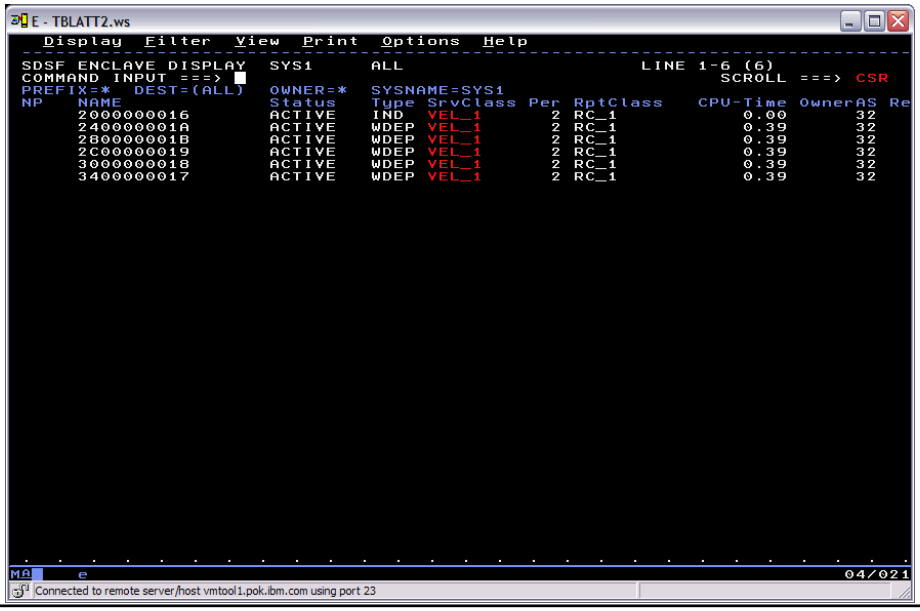
DB2 parallelism, WLM, and zIIPs

- DB2 Parallelism and zIIPs
  - Controlled by a CPU threshold. Once the threshold is met all child tasks are zIIP eligible
  - Parents are not zIIP eligible
  - Parent and child CPU time contribute to the CPU Threshold
  - Can see any kind of work, CICS, IMS, TSO, batch using zIIP resources
- DB2 will use new Work-Dependent Enclaves for Child tasks
  - APAR OA26104 for releases 1.8 and beyond
  - Without new Work Dependent Enclave support parallel enclaves must be classified using subsystem DB2
    - Unclassified work would wind up in SYSOTHER

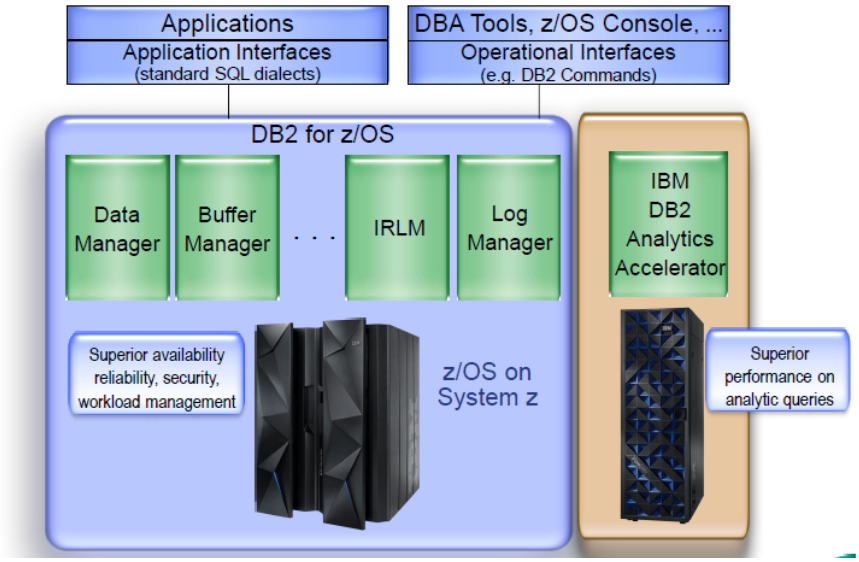
DDF and work-dependent enclaves



Work-dependent enclaves in SDSF



IBM DB2 Analytics Accelerator

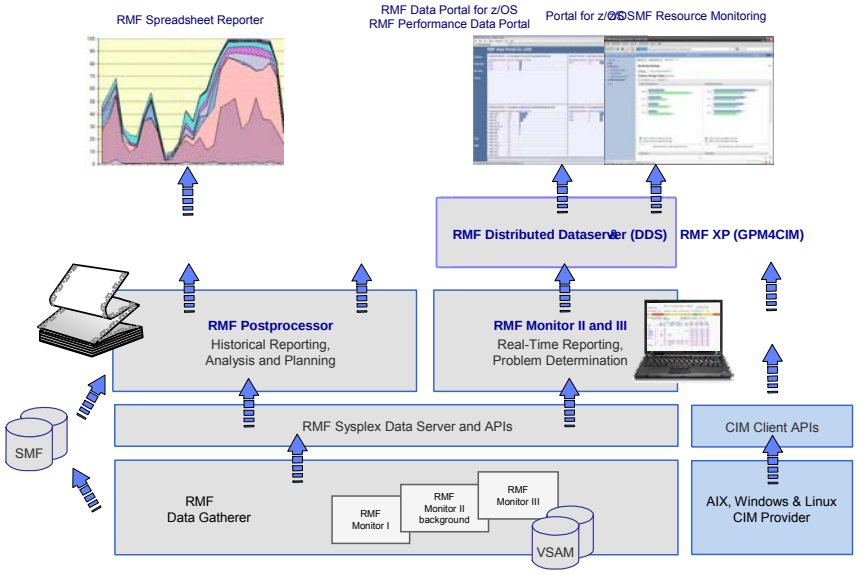


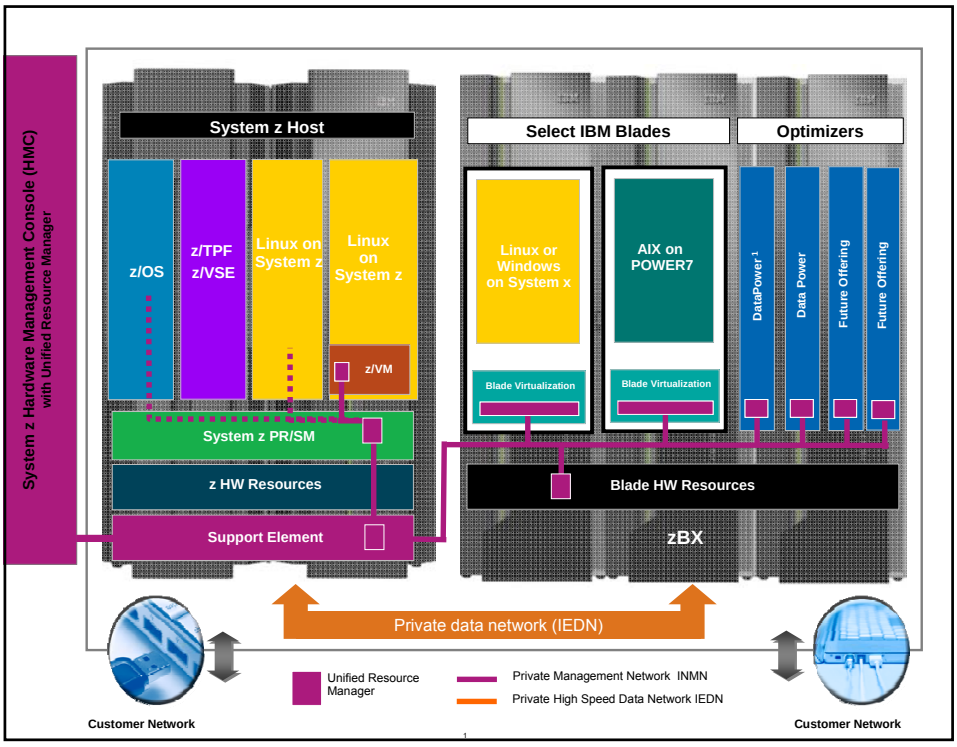
WLM and IDAA Interaction

- Workload Manager integration introduced in Version 3.1
  - DB2 detects WLM service class and importance level and sends it to the accelerator with each query submitted from a remote application.
  - The local applications such as SPUFI, TEP3, CICS, IMS are not supported
- The accelerator maps the importance level to a Netezza priority and alters the session prior to query execution, using the corresponding priority. Also threads scheduled will have their priorities adjusted.
- Version 4.1 extends the support to the local applications as well
- Mapping changes – apply to both remote and local applications

WLM Importance Level	Netezza Priority	
	Version 3	Version 4
System	Critical	Critical
Importance 1	Critical	Critical
Importance 2	High	Critical
Importance 3	Normal	High
Importance 4	Normal	Normal
Importance 5	Normal	Low
Discretionary	Low	Low

RMF Product Overview and RMF XP





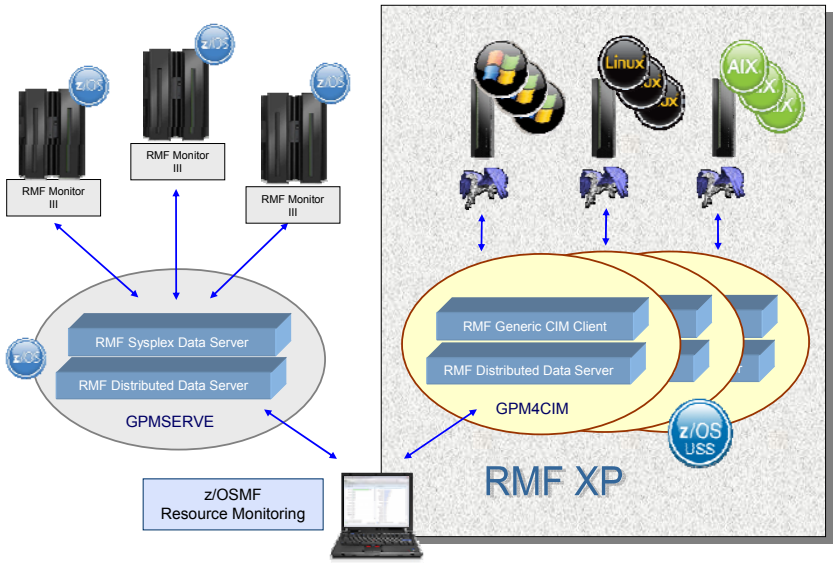
RMF XP Enhancements

IBM

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

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

RMF XP – Component Overview

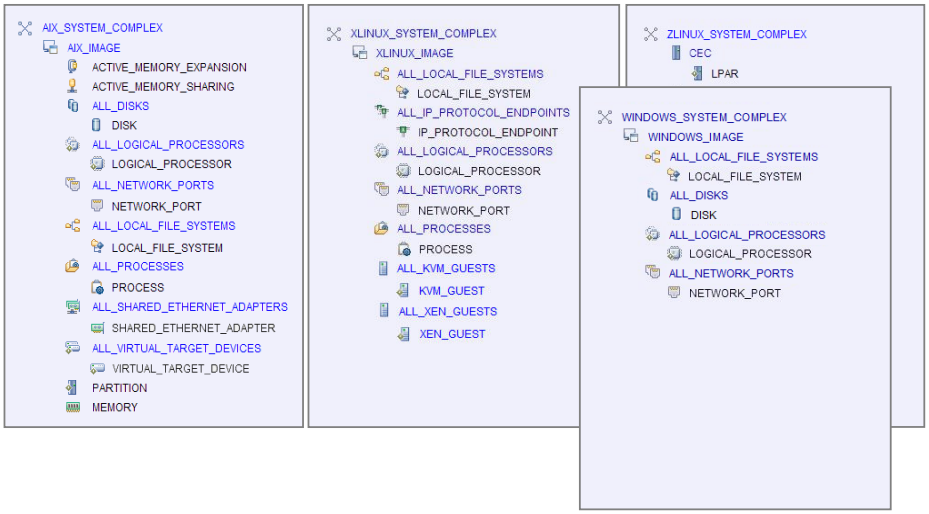


RMF XP Windows Support - Invocation

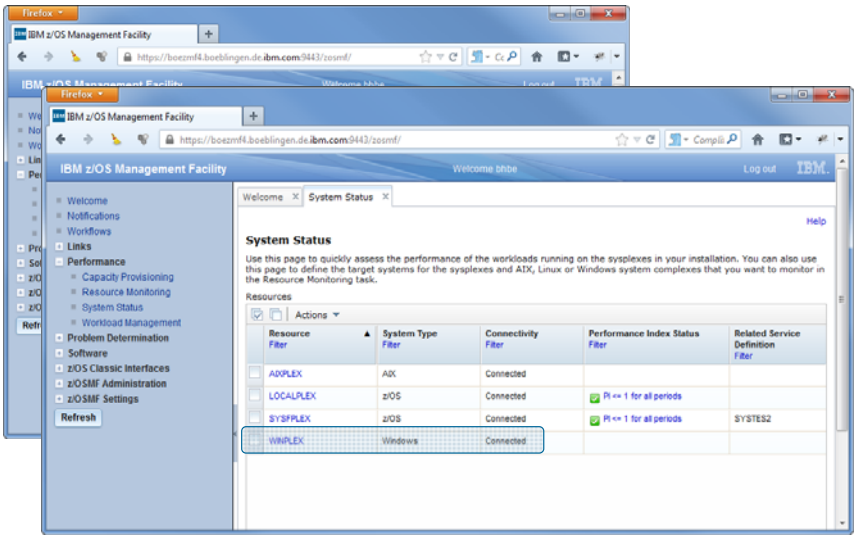
- Started Task: SYS1.PROCLIB(GPM4CIM)
- Runs in USS Environment via BPXBATCH
- Multiple instances can run in parallel: one STC per platform
  - S GPM4CIM.GPM4A,OS=A
  - S GPM4CIM.GPM4X,OS=X
  - S GPM4CIM.GPM4Z,OS=Z
  - S GPM4CIM.GPM4W,OS=W

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

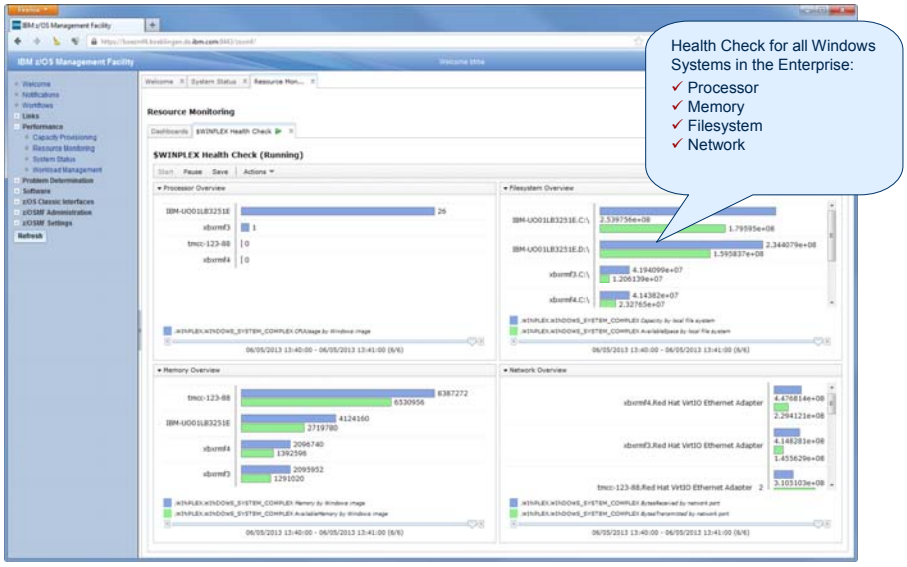
RMF XP Windows Support – Resource Model



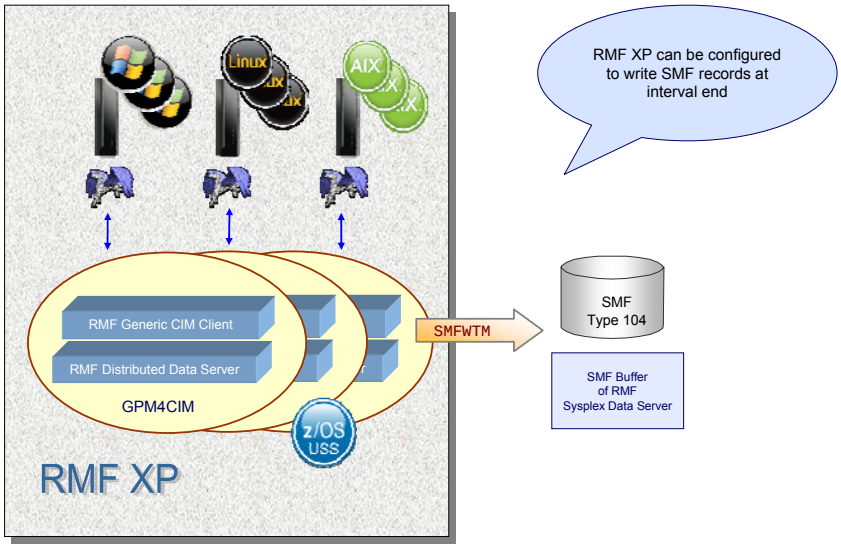
RMF XP Windows Support – z/OSMF Resource Monitoring



RMF XP Windows Support – z/OSMF Resource Monitoring



RMF XP & SMF Records



RMF XP & SMF Records

One Subtype per Metric Category

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

RMF XP & SMF Records

One Subtype per Metric Category

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

Blocked Workload Support: RMF

```
                                C P U   A C T I V I T Y
…

BLOCKED WORKLOAD ANALYSIS

OPT PARAMETERS: BLWLTRPCT (%)   0.5   PROMOTE RATE:  DEFINED    50000   WAITERS FOR PROMOTE:  AVG   0.001
                BLWLINTHD        60                 USED (%)      95                         PEAK     15
```

- **Extensions of RMF Postprocessor CPU Activity and WLMGL reports with information about blocked workloads and the temporary promotion of their dispatching priority**
- **SMF record 70-1 (CPU activity) and SMF 72-3 (Workload activity)**

Promoted transactions: RMF workload activity report

```
                                W O R K L O A D   A C T I V I T Y
                                                                                                  PAGE   1
   z/OS V1R13              SYSPLEX SVPLEX3           DATE 09/28/2011          INTERVAL 15.00.003   MODE = GOAL
                           RPT VERSION V1R13 RMF     TIME 17.00.00

                                   POLICY ACTIVATION DATE/TIME 09/14/2011 11.08.09

------------------------------------------------------------------------------------------------------------------- SERVICE CLASS(ES)

 REPORT BY: POLICY=BASEPOL    WORKLOAD=STC_WLD    SERVICE CLASS=STCLOW     RESOURCE GROUP=*NONE
                                                  CRITICAL     =NONE
                                                  DESCRIPTION  =Low priority for STC workloads

 -TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---  --PROMOTED--  ----STORAGE----
 AVG     153.37  ACTUAL         3.02.885  SSCHRT  56.9  IOC      3964   CPU  805,697  CP     92.24  BLK    1.489  AVG     1195.43
 MPL     152.35  EXECUTION      3.02.391  RESP    15.1  CPU     15184K  SRB   13.850  AAPCP   0.00  ENQ    0.046  TOTAL  182122.4
 ENDED      599  QUEUED              494  CONN     1.3  MSO        0    RCT    0.995  IIPCP   0.00  CRM    5.593  SHARED   230.59
 END/S     0.67  R/S AFFIN             0  DISC     0.3  SRB   261005    IIT    0.576                LCK    0.000
 #SWAPS    3391  INELIGIBLE            0  Q+PEND   4.5  TOT    15449K   HST    0.000  AAP     0.00  SUP    0.000  -PAGE-IN RATES-
 EXCTD        0  CONVERSION        5.188  IOSQ     9.0  /SEC   17202    AAP    0.000  IIP     0.00                SINGLE      0.0
 AVG ENC   0.00  STD DEV        3.27.429                                 IIP    0.000                              BLOCK       0.0
 REM ENC   0.00                                        ABSRPTN   113                                             SHARED      0.0
 MS ENC    0.00                                        TRX SERV  112                                             HSP         0.0

------------------------------------------SERVICE CLASSES BEING SERVED-------------------------------------------------------------
 DB2LOW
```

IBM

Promoted transactions RMF workload activity report

```
SERVICE TIME  ---APPL %---  --PROMOTED--  ----STORAGE----
CPU  805.697  CP     92.24  BLK    1.489  AVG     1195.43
SRB   13.850  AAPCP   0.00  ENQ    0.046  TOTAL  182122.4
RCT    9.995  IIPCP   0.00  CRM    5.593  SHARED   230.59
IIT    0.576                LCK    0.000
HST    0.000  AAP     0.00  SUP    0.000  -PAGE-IN RATES-
AAP    0.000  IIP     0.00                  SINGLE      0.0
IIP    0.000                                BLOCK       0.0
                                            SHARED      0.0
                                            HSP         0.0

RVED-----------------------------------------------------
```

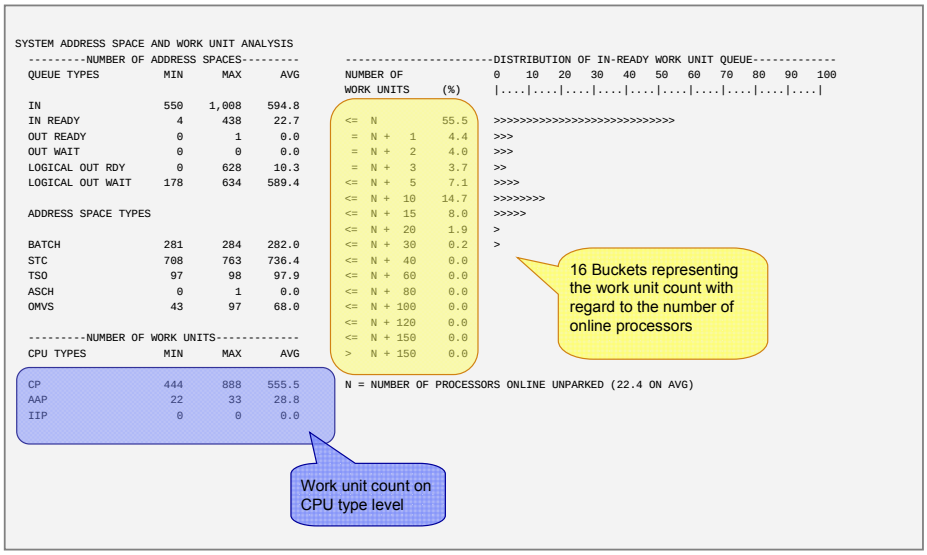
IBM

Promoted transactions RMF field definitions

CPU time in seconds that transactions in this group were running at a promoted dispatching priority, separated by the reason for the promotion:

BLK CPU time in seconds consumed while the dispatching priority of work with low importance was temporarily raised to help blocked workloads
ENQ CPU time in seconds consumed while the dispatching priority was temporarily raised by enqueue management because the work held a resource that other work needed.
CRM CPU time in seconds consumed while the dispatching priority was temporarily raised by chronic resource contention management because the work held a resource that other work needed
LCK In HiperDispatch mode, the CPU time in seconds consumed while the dispatching priority was temporarily raised to shorten the lock hold time of a local suspend lock held by the work unit.
SUP CPU time in seconds consumed while the dispatching priority for a work unit was temporarily raised by the z/OS supervisor to a higher dispatching priority than assigned by WLM.

Work unit queue distribution: Mon I CPU report



```
SYSTEM ADDRESS SPACE AND WORK UNIT ANALYSIS
  ---------NUMBER OF ADDRESS SPACES---------      ------------------------DISTRIBUTION OF IN-READY WORK UNIT QUEUE-------------
  QUEUE TYPES           MIN       MAX      AVG    NUMBER OF              0    10   20   30   40   50   60   70   80   90   100
                                                  WORK UNITS     (%)     |....|....|....|....|....|....|....|....|....|....|
  IN                    550     1,008    594.8
  IN READY                4       438     22.7    <=  N           55.5   >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
  OUT READY               0         1      0.0     =  N +   1     4.4   >>>
  OUT WAIT                0         0      0.0     =  N +   2     4.0   >>>
  LOGICAL OUT RDY         0       628     10.3     =  N +   3     3.7   >>
  LOGICAL OUT WAIT      178       634    589.4    <=  N +   5     7.1   >>>>
                                                  <=  N +  10    14.7   >>>>>>>>
  ADDRESS SPACE TYPES                             <=  N +  15     8.0   >>>>>
                                                  <=  N +  20     1.9   >
  BATCH                 281       284    282.0    <=  N +  30     0.2   >
  STC                   708       763    736.4    <=  N +  40     0.0
  TSO                    97        98     97.9    <=  N +  60     0.0
  ASCH                    0         1      0.0    <=  N +  80     0.0
  OMVS                   43        97     68.0    <=  N + 100     0.0
                                                  <=  N + 120     0.0
  ---------NUMBER OF WORK UNITS-------------       <=  N + 150     0.0
  CPU TYPES             MIN       MAX      AVG    >   N + 150     0.0

  CP                    444       888    555.5    N = NUMBER OF PROCESSORS ONLINE UNPARKED (22.4 ON AVG)
  AAP                    22        33     28.8
  IIP                     0         0      0.0
```

16 Buckets representing the work unit count with regard to the number of online processors

Work unit count on CPU type level

IBM Notice Regarding Specialty Engines

Any information contained in this document regarding Specialty Engines ("SEs") and SE eligible workloads provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs). IBM authorizes customers to use IBM SEs only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at:
www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT").

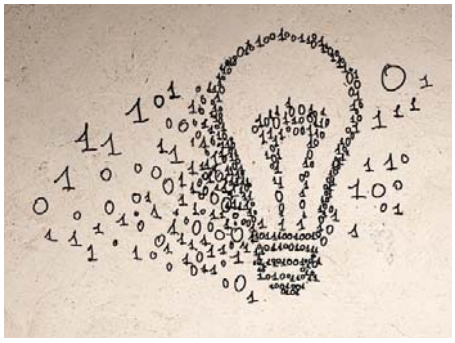
No other workload processing is authorized for execution on an SE.

IBM offers SEs at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

Glenn Anderson, IBM Lab Services and Training

Top New z/OS Performance Functions Every Sysprog Should Understand

Thank you for attending!



Session 15219