Top New z/OS Performance Functions Every Sysprog Should Understand

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 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))
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
Record Program

- 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

- CPU Reduction due to compression performed by zEDC
- With compatibility PTFs, software inflate can be done on downlevel z/OS or processor without zEDC
- Logger storage requirements minimized due to compression
- Fewer offloads should result in CPU savings
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
    - IQPPFMBPCIE 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

![Diagram showing access time for different memory tiers]

- CPU: < 20 ns
- Cache: < 200 ns
- Random Access Memory (RAM): 5-20 micro sec.
- Flash Express: 1-3 ms
- Solid State Drive (SSD) Storage: < 10 ms
- Spinning Disk Drive: < seconds
- WORM, Tape Library: seconds
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

• 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

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
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

---

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.
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.)

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 use flash for paging
  - ALL is the default
    - Default indicates all flash defined to the partition is available for paging
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
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 & Pageable Large Pages RMF Support**

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

```
<table>
<thead>
<tr>
<th></th>
<th>SHARED FRAMES</th>
<th>TOTAL SLOTS</th>
<th>CENTRAL STORAGE</th>
<th>FIXED TOT</th>
<th>FIXED BEL</th>
<th>AUX DASD</th>
<th>AUX SCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>7,937</td>
<td>44</td>
<td>30</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>7,937</td>
<td>44</td>
<td>30</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AVG</td>
<td>7,937</td>
<td>44</td>
<td>30</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>LOCAL PAGE DATA SET SLOTS</th>
<th>TOTAL</th>
<th>AVAILABLE</th>
<th>BAD</th>
<th>NON-VIO</th>
<th>VI O</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>5,399,997</td>
<td>4,269,302</td>
<td>0</td>
<td>1,128,251</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>5,399,997</td>
<td>4,271,746</td>
<td>0</td>
<td>1,130,695</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AVG</td>
<td>5,399,997</td>
<td>4,270,524</td>
<td>0</td>
<td>1,130,473</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SCM PAGING BLOCKS</th>
<th>TOTAL</th>
<th>AVAILABLE</th>
<th>BAD</th>
<th>IN-USE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MAX</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AVG</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
```

The number of shared pages backed on SCM

System wide statistics of 4K SCM paging blocks as: Total, Available, Unavailable and Used 4K blocks
zEC12 – Flash Memory & Pageable Large Pages...

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

**PAGING ACTIVITY**

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

<table>
<thead>
<tr>
<th>MEMORY OBJECTS</th>
<th>COMMON</th>
<th>SHARED</th>
<th>1 MB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MIN</strong></td>
<td>93</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>MAX</strong></td>
<td>56</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><strong>AVG</strong></td>
<td>54</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LMB FRAMES</th>
<th>TOTAL</th>
<th>FIXED</th>
<th>PAGEABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MIN</strong></td>
<td>200</td>
<td>80</td>
<td>120</td>
</tr>
<tr>
<td><strong>MAX</strong></td>
<td>200</td>
<td>170</td>
<td>120</td>
</tr>
<tr>
<td><strong>AVG</strong></td>
<td>200</td>
<td>136</td>
<td>64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HIGH SHARED FRAMES</th>
<th>TOTAL</th>
<th>CENTRAL STORAGE</th>
<th>AUX DASD</th>
<th>AUX SCM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MIN</strong></td>
<td>136902.1M</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>MAX</strong></td>
<td>136902.1M</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>AVG</strong></td>
<td>136902.1M</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HIGH COMMON FRAMES</th>
<th>TOTAL</th>
<th>CENTRAL STORAGE</th>
<th>FIXED 4K</th>
<th>AUX DASD</th>
<th>AUX SCM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MIN</strong></td>
<td>17301504</td>
<td>6,090</td>
<td>2,773</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>MAX</strong></td>
<td>17301504</td>
<td>6,093</td>
<td>2,733</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>AVG</strong></td>
<td>17301504</td>
<td>6,092</td>
<td>2,733</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**zEC12 – Flash Memory & Pageable Large Pages...**

- New SCM statistics RMF Postprocessor Page Data Set Activity report

**PAGE DATA SET ACTIVITY**

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

<table>
<thead>
<tr>
<th>PAGE TYPE</th>
<th>SERIAL NUM</th>
<th>DEVICE</th>
<th>SLots</th>
<th>SLots Used</th>
<th>BAD</th>
<th>IN</th>
<th>TRAN</th>
<th>NUMBER</th>
<th>PAGES</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLPA</td>
<td>D406 33903</td>
<td>71999</td>
<td>16851</td>
<td>16851</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COMMON</td>
<td>D406 33903</td>
<td>35999</td>
<td>34</td>
<td>34</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LOCAL</td>
<td>D406 33903</td>
<td>593999</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**System wide statistics of 4K SCM paging block usage and SCM paging activity**
Statistics for Fixed Large Memory Objects in RMF Postprocessor

Virtual Storage Activity report

Decomposition into Fixed and Pageable Large Pages

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

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

New Large Page statistics in RMF Monitor III STORM report
Warning track

- 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

Logical Processor

Grace Period (50 μs)

Physical Processor

Processor Running Time: 12.5ms (*)

Warning Track Interruption

Has Z/OS time to return the LP to PR/SM?

Z/OS (re) Dispatch on a new LP then PR/SM (re) Dispatch on another PP

(*) Processor running time = 35ms * (PRT/PP)
12.5ms = PRT = 25ms

Latent Demand: LPAR Busy vs MVS Busy

<table>
<thead>
<tr>
<th>CPU</th>
<th>2097</th>
<th>CPC CAPACITY</th>
<th>1451</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL</td>
<td>719</td>
<td>GREASE REASONS</td>
<td>MVA</td>
</tr>
<tr>
<td>CSP</td>
<td>LPAR BUSY</td>
<td>MVS BUSY</td>
<td>PSS</td>
</tr>
<tr>
<td>NUM</td>
<td>SYSTEM</td>
<td>TIME</td>
<td>PSS</td>
</tr>
<tr>
<td>0 CP</td>
<td>100.00</td>
<td>96.77</td>
<td>96.00</td>
</tr>
<tr>
<td>1 CP</td>
<td>100.00</td>
<td>94.91</td>
<td>94.95</td>
</tr>
<tr>
<td>2 CP</td>
<td>100.00</td>
<td>94.72</td>
<td>94.74</td>
</tr>
<tr>
<td>3 CP</td>
<td>100.00</td>
<td>95.07</td>
<td>95.10</td>
</tr>
<tr>
<td>4 CP</td>
<td>100.00</td>
<td>50.18</td>
<td>93.55</td>
</tr>
<tr>
<td>5 CP</td>
<td>100.00</td>
<td>50.15</td>
<td>93.56</td>
</tr>
<tr>
<td>6 CP</td>
<td>100.00</td>
<td>50.10</td>
<td>93.59</td>
</tr>
<tr>
<td>7 CP</td>
<td>100.00</td>
<td>51.40</td>
<td>80.16</td>
</tr>
<tr>
<td>8 CP</td>
<td>100.00</td>
<td>52.12</td>
<td>81.69</td>
</tr>
<tr>
<td>9 CP</td>
<td>100.00</td>
<td>56.12</td>
<td>87.87</td>
</tr>
<tr>
<td>A CP</td>
<td>100.00</td>
<td>48.37</td>
<td>86.74</td>
</tr>
<tr>
<td>B CP</td>
<td>100.00</td>
<td>38.46</td>
<td>06.76</td>
</tr>
<tr>
<td>C CP</td>
<td>100.00</td>
<td>38.00</td>
<td>06.98</td>
</tr>
<tr>
<td>D CP</td>
<td>100.00</td>
<td>18.20</td>
<td>08.13</td>
</tr>
<tr>
<td>E CP</td>
<td>100.00</td>
<td>0.00</td>
<td>----</td>
</tr>
<tr>
<td>F CP</td>
<td>100.00</td>
<td>0.00</td>
<td>----</td>
</tr>
<tr>
<td>10 CP</td>
<td>100.00</td>
<td>0.00</td>
<td>----</td>
</tr>
<tr>
<td>TOTAL/AVG</td>
<td>42.47</td>
<td>91.45</td>
<td>539.0</td>
</tr>
</tbody>
</table>

CEC Busy = 98.85

0.0115 * 19 CP = 22 CPs available

Weight: 5.32 CPs

Using: 42.47/100 * 17

LCP = 7.22 CPs

(c) 2014 IBM Corporation
Warning track statistics

- 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.

<table>
<thead>
<tr>
<th>Offset</th>
<th>Name</th>
<th>Length</th>
<th>Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 x50</td>
<td>SMF70WTS</td>
<td>4</td>
<td>Binary</td>
<td>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.</td>
</tr>
<tr>
<td>84 x54</td>
<td>SMF70WUU</td>
<td>4</td>
<td>Binary</td>
<td>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.</td>
</tr>
<tr>
<td>88 x58</td>
<td>SMF70WTTI</td>
<td>4</td>
<td>Binary</td>
<td>Amount of time in microseconds that a logical processor was yielded to PR/SM due to Warning Track processing.</td>
</tr>
</tbody>
</table>

RMF Postprocessor Overview Conditions

<table>
<thead>
<tr>
<th>Name</th>
<th>Qualifier</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTRKCP</td>
<td>cpu-id</td>
<td>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.</td>
</tr>
<tr>
<td>WTRKTPC</td>
<td>cpu-id</td>
<td>Time in microseconds that a processor was yielded to PR/SM due to Warning Track processing.</td>
</tr>
</tbody>
</table>

Work-dependent enclaves in SDSF
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 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

In cases where DRDA applications create extended duration work threads in DB2, for example through extensive use of held cursors, the zIIP utilization levels can become more variable. DB2 and DDF now may use **work-dependent enclaves** in this situation to control this variability. See APAR PM28626.
Work-dependent enclaves in SDSF

IBM DB2 Analytics Accelerator

Applications
Application Interfaces
(standard SQL dialects)

DB2 for z/OS
- Data Manager
- Buffer Manager
- IRLM
- Log Manager

DBA Tools, z/OS Console,
Operational Interfaces
(e.g. DB2 Commands)

IBM DB2 Analytics Accelerator

Superior availability, reliability, security, workload management

z/OS on System z

Superior performance on analytic queries
### 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.

<table>
<thead>
<tr>
<th>WLM Importance Level</th>
<th>Netezza Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Version 3</td>
</tr>
<tr>
<td>System</td>
<td>Critical</td>
</tr>
<tr>
<td>Importance 1</td>
<td>Critical</td>
</tr>
<tr>
<td>Importance 2</td>
<td>High</td>
</tr>
<tr>
<td>Importance 3</td>
<td>Normal</td>
</tr>
<tr>
<td>Importance 4</td>
<td>Normal</td>
</tr>
<tr>
<td>Importance 5</td>
<td>Normal</td>
</tr>
<tr>
<td>Discretionary</td>
<td>Low</td>
</tr>
</tbody>
</table>

### RMF Product Overview and RMF XP

- **RMF Monitor II and III**
  - Real-Time Reporting, Problem Determination
- **RMF Monitor I**
  - Background
- **SMF**
  - Data Gatherer
- **RMF Postprocessor**
  - Historical Reporting, Analysis and Planning
- **RMF Spreadsheet Reporter**
- **RMF Distributed Dataserver (DDS)**
- **RMF Data Portal for z/OS**
- **RMF Performance Data Portal**
- **Portal for z/OS SMF Resource Monitoring**
- **RMF XP (GPM4CIM)**
- **CIM Client APIs**
- **AIX, Windows & Linux CIM Provider**
RMF XP Enhancements

- RMF XP is the solution for Cross Platform Performance Monitoring
- RMF XP supports the Operating Systems running on
  - x Blades
  - p Blades
- In addition RMF XP supports Linux on System z
  - LPAR Mode
  - VM Guest Mode
RMF XP – Component Overview

RMF Generic CIM Client
RMF Distributed Data Server
GPM4CIM
RMF Generic CIM Client
RMF Distributed Data Server
GPM4CIM
z/OSMF
Resource Monitoring

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

```bash
//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=*  
//SYSTOUT DD   SYSOUT=*  
```

(c) 2014 IBM Corporation
RMF XP Windows Support – Resource Model

RMF XP Windows Support – z/OSMF Resource Monitoring
RMF XP Windows Support – z/OSMF Resource Monitoring

Health Check for all Windows Systems in the Enterprise:
- Processor
- Memory
- Filesystem
- Network

RMF XP & SMF Records

RMF XP can be configured to write SMF records at interval end
## RMF XP & SMF Records

### Linux on System z

<table>
<thead>
<tr>
<th>Metric Category</th>
<th>ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIX on System p</td>
<td>1</td>
</tr>
<tr>
<td>AIX_Process</td>
<td>2</td>
</tr>
<tr>
<td>AIX_ComputerSystem</td>
<td>3</td>
</tr>
<tr>
<td>AIX_Disk</td>
<td>4</td>
</tr>
<tr>
<td>AIX_NetworkPort</td>
<td>5</td>
</tr>
<tr>
<td>AIX_FileSystem</td>
<td>6</td>
</tr>
<tr>
<td>AIX_Memory</td>
<td>7</td>
</tr>
<tr>
<td>AIX_OperatingSystem</td>
<td>8</td>
</tr>
<tr>
<td>AIX_Process</td>
<td>9</td>
</tr>
<tr>
<td>AIX_SharedEthernetAdapter</td>
<td>10</td>
</tr>
<tr>
<td>AIX_ActiveMemorySharing</td>
<td>11</td>
</tr>
<tr>
<td>AIX_VirtualTargetDevice</td>
<td>12</td>
</tr>
</tbody>
</table>

### Linux on System x

<table>
<thead>
<tr>
<th>Metric Category</th>
<th>ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIX on System x</td>
<td>ST</td>
</tr>
<tr>
<td>AIX_Process</td>
<td>20</td>
</tr>
<tr>
<td>AIX_ComputerSystem</td>
<td>21</td>
</tr>
<tr>
<td>AIX_Disk</td>
<td>22</td>
</tr>
<tr>
<td>AIX_NetworkPort</td>
<td>24</td>
</tr>
<tr>
<td>AIX_FileSystem</td>
<td>25</td>
</tr>
<tr>
<td>AIX_Memory</td>
<td>26</td>
</tr>
<tr>
<td>AIX_OperatingSystem</td>
<td>27</td>
</tr>
<tr>
<td>AIX_Process</td>
<td>28</td>
</tr>
<tr>
<td>AIX_SharedEthernetAdapter</td>
<td>29</td>
</tr>
<tr>
<td>AIX_ActiveMemorySharing</td>
<td>30</td>
</tr>
<tr>
<td>AIX_VirtualTargetDevice</td>
<td>31</td>
</tr>
</tbody>
</table>

### Linux on System z

<table>
<thead>
<tr>
<th>Metric Category</th>
<th>ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linux on System z</td>
<td>ST</td>
</tr>
<tr>
<td>Linux_PPProtocolEndpoint</td>
<td>20</td>
</tr>
<tr>
<td>Linux_LocalFileSystem</td>
<td>21</td>
</tr>
<tr>
<td>Linux_NetworkPort</td>
<td>22</td>
</tr>
<tr>
<td>Linux_OperatingSystem</td>
<td>23</td>
</tr>
<tr>
<td>Linux_Processor</td>
<td>24</td>
</tr>
<tr>
<td>Linux_LocalFileSystem</td>
<td>25</td>
</tr>
<tr>
<td>Linux_NetworkPort</td>
<td>26</td>
</tr>
<tr>
<td>Linux_OperatingSystem</td>
<td>27</td>
</tr>
<tr>
<td>Linux_Processor</td>
<td>28</td>
</tr>
<tr>
<td>Linux_PPProtocolEndpoint</td>
<td>30</td>
</tr>
<tr>
<td>Linux_LocalFileSystem</td>
<td>31</td>
</tr>
</tbody>
</table>

### Linux on System x

<table>
<thead>
<tr>
<th>Metric Category</th>
<th>ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linux on System x</td>
<td>ST</td>
</tr>
<tr>
<td>Linux_PPProtocolEndpoint</td>
<td>60</td>
</tr>
<tr>
<td>Linux_NetworkPort</td>
<td>61</td>
</tr>
<tr>
<td>Linux_OperatingSystem</td>
<td>62</td>
</tr>
<tr>
<td>Linux_Processor</td>
<td>63</td>
</tr>
<tr>
<td>Linux_NetworkPort</td>
<td>64</td>
</tr>
</tbody>
</table>

### One Subtype per Metric Category

(c) 2014 IBM Corporation
Blocked Workload Support: RMF

- 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
Promoted transactions RMF workload activity report

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

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

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., zILPs, 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:


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.
Top New z/OS Performance Functions Every Sysprog Should Understand

Thank you for attending!