#SHAREorg

SHARE
Technology · Connections · Results

DFSMS Basics: VSAM

VSAM RLS Best Practices

How to leverage VSAM RLS best practices and performance in your environment

David LeGendre, dlegendr@us.ibm.com

Session : 10967

Trademarks

DFSMSdfp, DFSMSdss, DFSMShsm, DFSMSrmm, IBM, IMS, MVS, MVS/DFP, MVS/ESA, MVS/SP, MVS/XA,

OS/390, SANergy, and SP are trademarks of International Business Machines Corporation in the United States, other countries, or both.

AIX, CICS, DB2, DFSMS/MVS, Parallel Sysplex, OS/390, S/390, Seascape, and z/OS are registered trademarks of International Business Machines Corporation in the United States, other countries, or both.

Domino, Lotus, Lotus Notes, Notes, and SmartSuite are trademarks or registered trademarks of Lotus Development Corporation. Tivoli, TME, Tivoli Enterprise are trademarks of Tivoli Systems Inc. in the United States and/or other countries.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both. UNIX is a registered trademark in the United States and other countries licensed exclusively through The Open Group.

Other company, product, and service names may be trademarks or service marks of others.

Notices & Disclaimers

Product information and data has been reviewed for accuracy as of the date of initial publication. Product information and data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or programs(s) described herein at any time without notice.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such such products, programs or services available in all countries in which IBM operates or does business. Consult your local IBM representative or IBM Business Partner for information about the product and services available in your area.

Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does notinfringe IBM's intellectually property rights, may be used instead. It is the user's responsibility to evaluate and verify the operation of any non-IBM product, program or service.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS"WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR INFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. IBM is not responsible for the performance or interoperability of any non-IBM products discussed herein.

Notices & Disclaimers

The performance data contained herein was obtained in a controlled, isolated environment. Actual results that may be obtained in other operating environments may vary significantly. While IBM has reviewed each item for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere.

The responsibility for use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's or user's ability to evaluate and integrate them into their operating environment. Customers or users attempting to adapt these techniques to their own environments do so at their own risk. IN NO EVENT SHALL IBM BE LIABLE FOR ANY DAMAGE ARISING FROM THE USE OF THIS INFORMATION, INCLUDING BUT NOT LIMITED TO,

LOSS OF DATA, BUSINESS INTERRUPTION, LOSS OF PROFIT OR LOSS OF OPPORTUNITY.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not necessarily tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Introduction

There are a number of settings, parameters and sizing considerations to be made when configuring VSAM RLS environments. Additionally, there are many performance indicators and that can help to tune those options to a specific installation. The goal of this presentation is to help provide some insights into those topics as well as to supply guidance when problems are encountered.

Agenda

- VSAM RLS I/O Path and Buffer Management Overview
  - RLS sub components
  - GET path example
  - LRU algorithm
- Performance and Tuning Parameters
  - Buffer Pools
  - PARMLIB
  - Cache Structures
  - Lock Structures
  - Data Class
  - Request Level Parameters
  - Data Set Considerations

Agenda

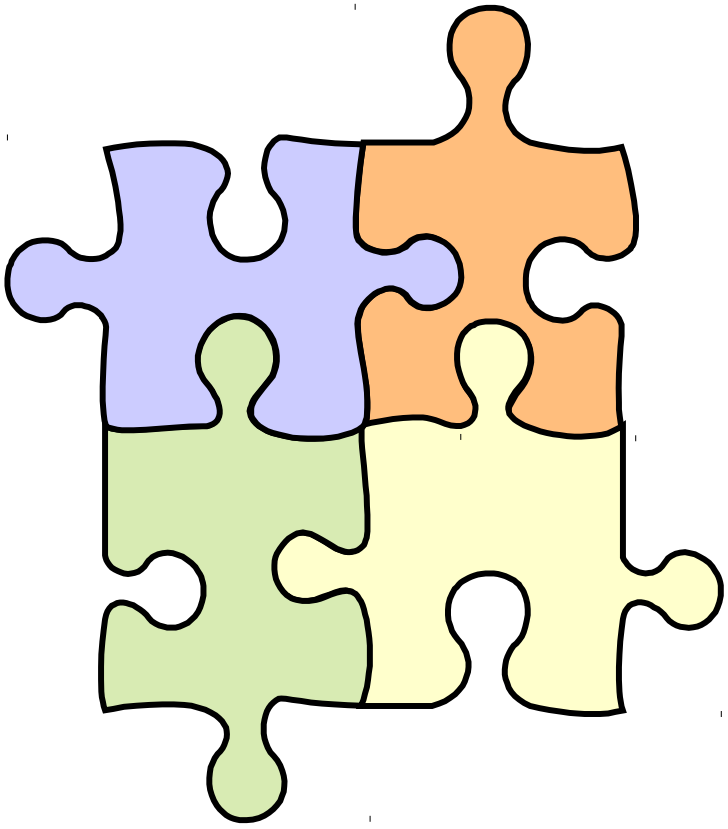
- Performance Measurements
  - SMF 64 records
  - SMF 42 subtypes 15-19
  - RMF Monitor III
  - RMF RLSSC & RLSDS
  - RMF RLSLRU
- SMSVSAM Diagnostics
  - Display commands
  - Hang / wait condition & SMSVSAM dump collection
  - Internal logic errors & required documentation

#SHAREorg

SHARE
Technology • Connections • Results

VSAM RLS I/O Path and Buffer Management Overview

VSAM RLS I/O Path – Sub Components



- VSAM Record Management (VRM)
  - Provides the VSAM interfaces: GET, PUT, POINT, ERASE, etc
  - Parameters passed to VRM are through the RPL control block
- Storage Management Locking Services (SMLS)
  - Interfaces with VRM and XCF locking services
  - Obtains, releases, and alters locks in the coupling facility lock structure(s)

VSAM RLS I/O Path – Sub Components

- Sysplex Cache Manager (SCM)
  - Interfaces with Buffer Management Facility and XCF caching services
  - Obtains directory elements and read/write data elements in the coupling facility cache structures

- Buffer Management Facility (BMF)
  - Interfaces with VRM and SCM to locate/add buffers to the local pool
  - Supports buffering past close
    - Data sets reopened for RLS within 10 minutes can reuse valid buffers remaining in the pool
  - Manages the size of the local buffer pool via a Least Recently Used (LRU) manager

VSAM RLS I/O Path – Sub Components

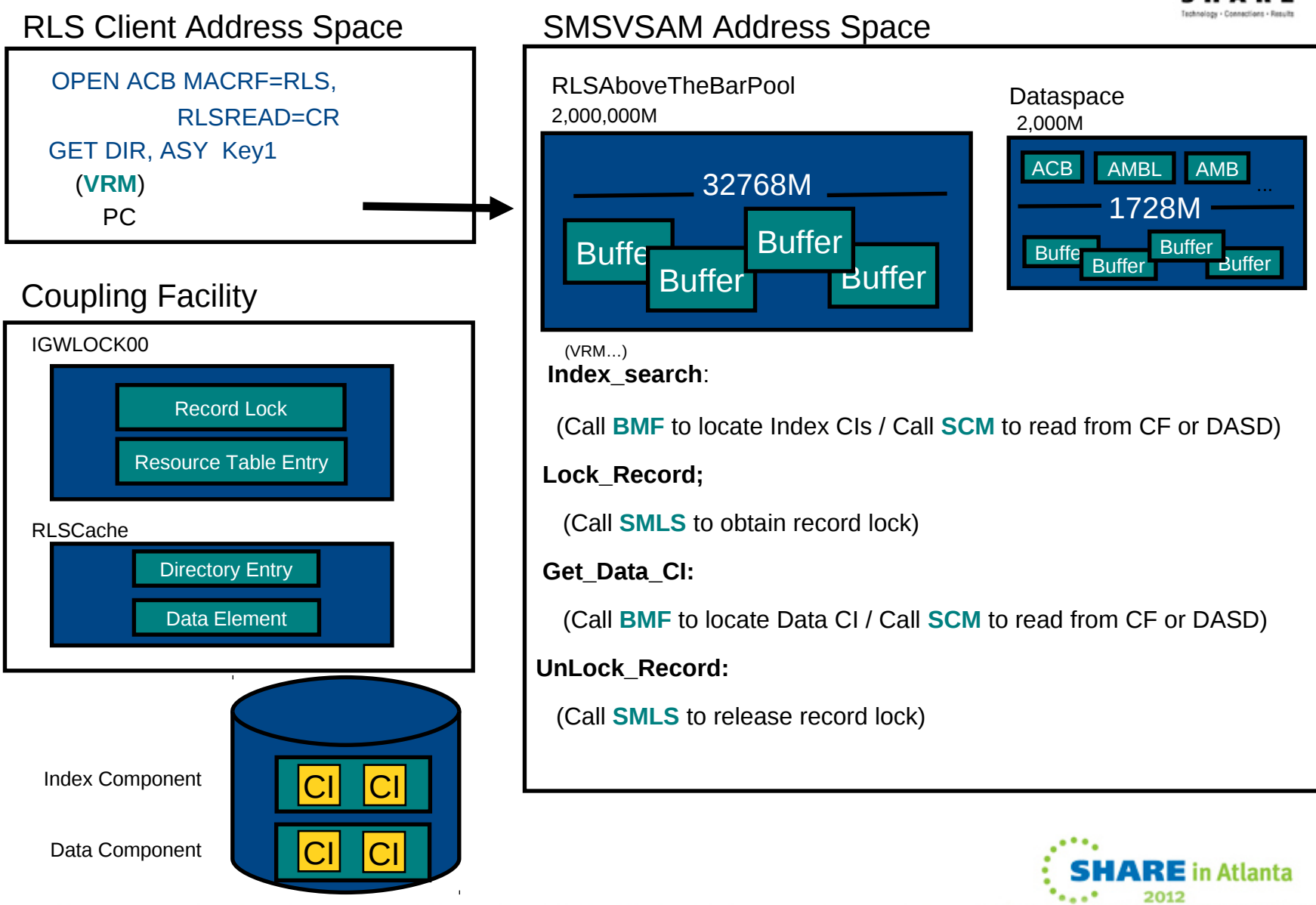
- **Performance Goal:** Spend the least amount of time in the I/O path!

GET/PUT ↔ VRM ↔ SMLS ↔ XCF Locking services

↔ BMF ↔ SCM ↔ XCF Caching Services

↔ Media Mgr Services (to DASD)

VSAM RLS I/O Path – Example of a GET

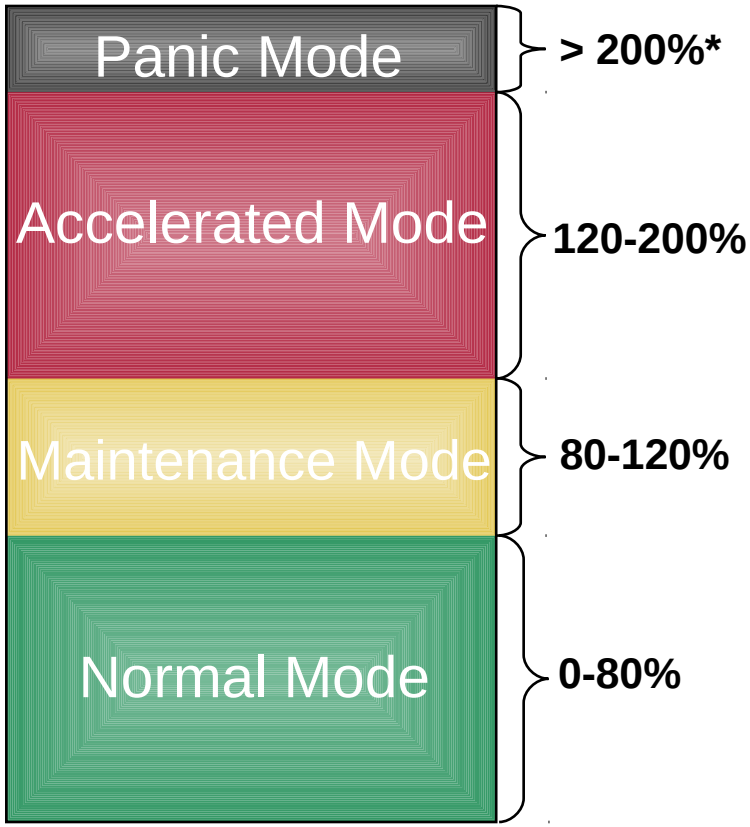


Buffer Management Overview

- 31 Bit Buffer Pool
  - Size: 10M to 1728M
  - Defined using the RLS_MAX_POOL_SIZE(nnnn) parameter
  - Same on all systems
- 64 Bit Buffer Pool
  - Size: 0 or 500M > size > 2,000,000M (2TB)
  - Defined using the RLSAboveTheBarMaxPoolSize(sysname1,nnnn) parameter
  - Can be specified per system

Buffer Management Overview – LRU(31 bit)

- 31 bit LRU Comprised of 4 different levels
  - **Normal Mode (0 – 80%)**
    - Max_Cycles_Unused = 240
    - Cycles_Unused = Cycles_Unused + 1
    - Cycle iterates every 15 seconds
    - (15 * 240) / 60 = 60 minutes
  - **Maintenance Mode (80 – 120%)**
    - Max_Cycles_Unused = Max_Cycles_Unused - 1
    - Cycles_Unused = Cycles_Unused + 1
    - Cycle iteration stays at 15 seconds
  - **Accelerated Mode (120 – 200%)**
    - Max_Cycles_Unused = Max_Cycles_Unused - 4
    - Cycles_Unused = Cycles_Unused + 1
    - Cycle iteration stays at 15 seconds
  - **Panic Mode (> 200%)**
    - Max_Cycles_Unused = Max_Cycles_Unused - 8
    - Cycles_Unused = Cycles_Unused + 1
    - Cycle iteration runs every 5 seconds

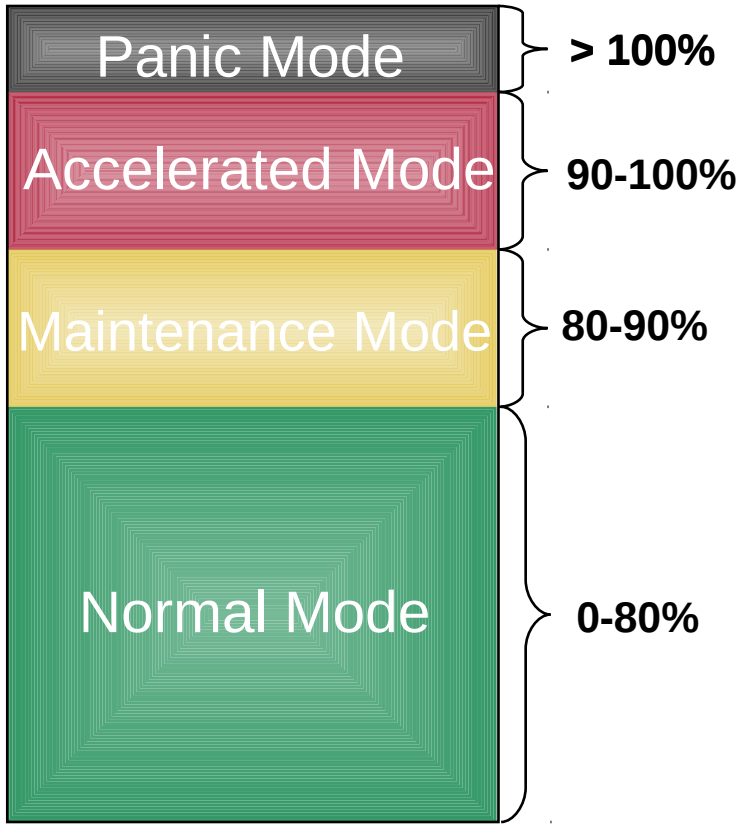


Percentage of RLS_MAX_POOL_SIZE in use

* *200% of RLS_MAX_POOL_SIZE OR > 1728MB Limit*

Buffer Management Overview – LRU(64 bit)

- 64 bit LRU Comprised of 4 different modes
  - **Normal Mode (0 - 80%)**
    - Buffers older than 60 minutes are recycled
  - **Maintenance Mode (80 - 90%)**
    - Buffers older than 60 minutes are recycled
  - **Accelerated Mode (90 - 100%)**
    - Buffers older than 30 minutes are recycled
    - Buffer stealing will be attempted
    - If no available buffers a new one will be generated
  - **Panic Mode (> 100%)**
    - Buffers older than 5 minutes are recycled
    - Buffer stealing will be attempted
    - If no available buffers a new one will be generated



Percentage of RLSAboveTheBarMaxPoolSize in use

Performance and Tuning Parameters

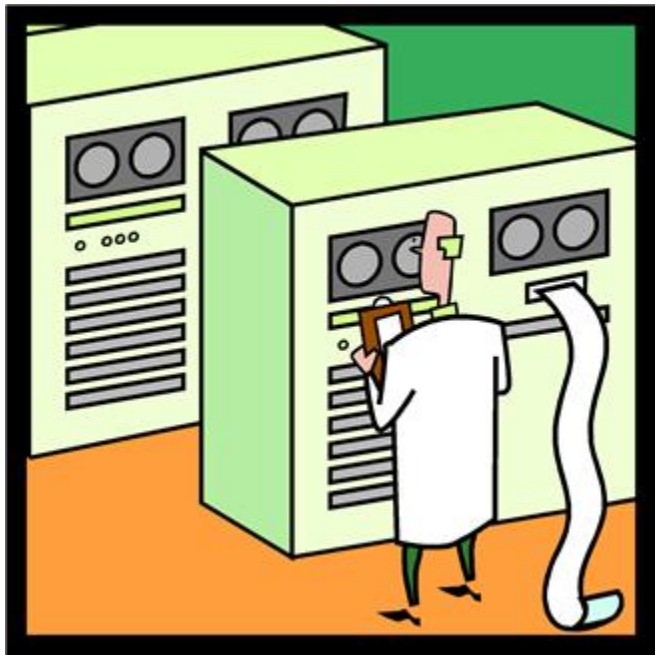
Performance and Tuning – Buffer Pools

- SYS1.PARMLIB(IGDSMSxx)
  - RLS_MAX_POOL_SIZE(nnnn MB)
    - Max size of SMSVSAM local buffer pool
    - nnnn = 10 to 9999, anything over 1500 is treated as the max of 1728
    - Recommendation: RLS_MAX_POOL_SIZE(<= 850)
      - *Must have adequate cache structure sizes*
      - *680M (80%) of buffers to reside in the 31 bit pool for one hour via LRU normal mode*
      - *Allows for a doubling of the pool to 1700M before entering panic mode (after exceeding the 1728M limit)*

Performance and Tuning – Buffer Pools

- SYS1.PARMLIB(IGDSMSxx)
  - RLSAboveTheBarMaxPoolSize([sys1,val1;sys2,val2...] | [ALL,value])
    - Amount of virtual storage allocated above 2G bar for RLS buffering
    - nnnn = 0, or 500M > size > 2,000,000M (2 TB)
    - Recommendation: RLSAboveTheBarMaxPoolSize(<= 32768)
      - *Must have matching amount of real storage*
      - *Must have adequate cache structure sizes*
      - *Recommended when LRU for 31 bit pool regularly in accelerated / panic mode*
      - *Great for applications needing to traverse large amounts of data repeatedly*

Performance and Tuning – Buffer Pools



- Pool Size values are a goal for which the LRU tries to maintain.
- If more buffers are required at any given time, the pool may temporarily exceed the values set.
- Total size of buffer pools should not exceed amount of real storage.
- A paged out buffer is immediately freed by the LRU.

Performance and Tuning – PARMLIB

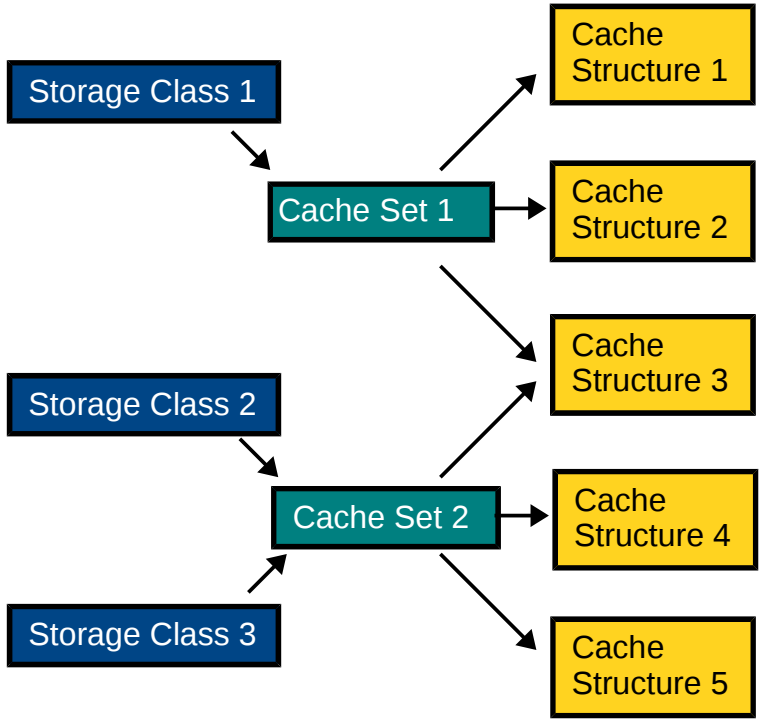
- SYS1.PARMLIB(IGDSMSxx)
  - RLSINIT(YES | NO)
    - SMSVSAM started during system init
    - Set this to YES or issue V SMS,SMSVSAM,ACTIVE at IPL
  - CF_TIME(xx seconds | 3600)
    - Interval for recording of SMF 42 subtype 15 – 19
    - The default is once / hour
    - The standard is generally once every 15 minutes
  - SMF_TIME(YES | NO)
    - Synchronizes creation of SMF 42 subtypes 2, 15 – 19 with RMF intervals
    - “YES” helps when correlating data between SMF and RMF during performance diagnosis

Performance and Tuning – PARMLIB

- SYS1.PARMLIB(IGDSMSxx)
  - RLS_MaxCFFeatureLevel(Z | A)
    - Dictates size of data placed in CF cache structures
    - Z: only CIs less than 4K cached
      - *Saves space in the cache structure*
      - *Advantageous if data is primarily read-only and remains valid in local buffers*
    - A: CIs up to 32K cached
      - *Requires more space in the cache structure*
      - *Advantageous when shared data is updated across the sysplex*
      - *Allows use of RLS CF Cache Feature in data class*
  - CA_RECLAIM (NONE | {DATACLAS | DATACLASS})
    - Enabling CA reclaim reduces KSDS fragmentation and the need for reorgs
    - None: CA reclaim is not used, regardless of data class setting
    - DATACLAS | DATACLASS: Enables CA reclaim for all eligible data sets

Performance and Tuning – Cache Structures

- Cache Structures and Cache Sets
  - Allows for partitioning of workloads and data sets
  - Structures and sets defined in the base config (share control data set)
  - Many cache structures mapped to one cache set
  - One cache set assigned to each storage class
  - Structure allocation details defined in CFRM policy
- Proper cache structure sizing is required to ensure maximum buffer effectiveness!



Performance and Tuning – Cache Structures

- The ideal theoretical *single* cache structure size:
  - Sum of 31 bit and 64 bit buffer goals
    - 31 bit buffer goal = ((RLS_Max_Pool_Size) * Number of Systems)
    - 64 bit buffer goal =
      (RLSAboveTheBarMaxPoolSize(system1) +
      RLSAboveTheBarMaxPoolSize(system2) + … +
      RLSAboveTheBarMaxPoolSize(systemN))
  - Assumptions:
    - RLS_MaxCFFeaturelevel(A): caching all data
    - No sharing of data across the sysplex

Performance and Tuning – Cache Structures

- Example:
  - RLS_Max_Pool_Size(850)
  - Number of Systems = 2
  - RLSAboveTheBarMaxPoolSize(System1,2048)
  - RLSAboveTheBarMaxPoolSize(System2,4096)
    - Cache_Structure_Size = (850*2) + 2048 + 4096 = 7844M
  - Cache structure sizes less than the ideal amount should be closely monitored for directory reclaims.

Performance and Tuning – Cache Structures

- Recommendations
  - Allocate different cache structures for different applications.
  - Allocate different structures to match internal structure usages.
- If more than one cache structure to be allocated:
  - Divide total ideal theoretical allocation across all structures
  - Base sizes on data set usage distribution
    - Size
    - Number
    - Amount of data accessed

Performance and Tuning – Cache Structures

- False Invalids:
  - Take place when a cache structure(s) too small for current buffer allocations
  - Cache structures invalidate entries to make room for new requests
  - Causes additional requests to DASD!
  - Tracked in
    - RMF panels RLSSC, and RLSDS
    - SMF 42 subtypes 15-16

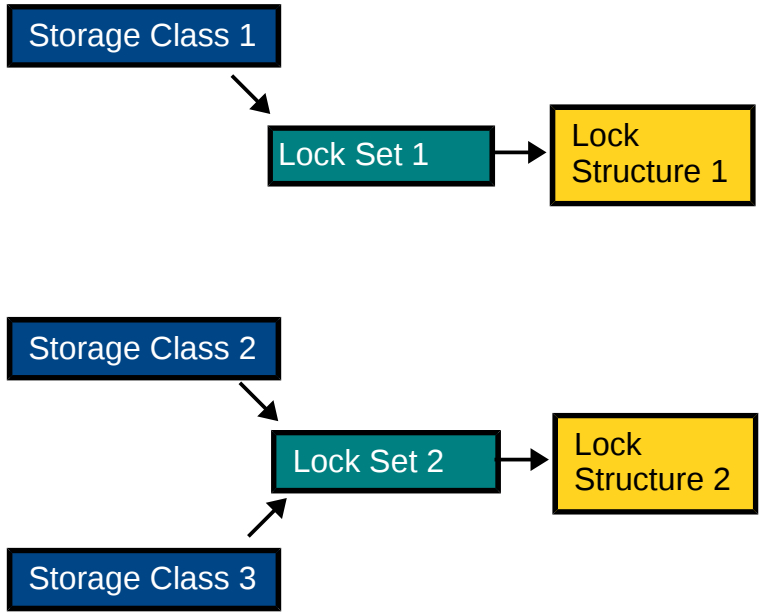
Performance and Tuning – Cache vs Buffer

- A balance between buffer and cache allocations must be made for optimum performance
- buffer space = large and cache space = small
  - Bottleneck in cache will occur between buffers and DASD
  - Cache will invalidate buffer entries to make room for new requests: thrashing
  - False Invalids occur when the cache structure needs more space than it has to satisfy CI requests.
- buffer space = small and cache space = large
  - Potentially wasted coupling facility space is expensive
  - Cache will try to keep as much data in it as possible
  - Buffer LRU will be the constraining factor

Performance and Tuning – Lock Structures

Lock Structures and Lock Sets

- Allows for partitioning workloads and environments (i.e. dev and prod)
- Structures and sets defined in the base config (SCDS)
- One lock structure mapped to one lock set
- One lock set assigned to each storage class
- Structure allocation details defined in CFRM policy



Performance and Tuning – Lock Structures

- Lock_Structure_Size =

  (10M * Number_Of_Systems * Lock_Entry_Size)

  - Lock_Entry_Size (depends on the CFRM MAXSYSTEM value)
    - MAXSYSTEM <= 7 Lock_Entry_Size = 2
    - MAXSYSTEM >= 8 & <24 Lock_Entry_Size = 4
    - MAXSYSTEM >=24 & <=32 Lock_Entry_Size = 8

- Example: MAXSYSTEM = 16 and 8 systems in sysplex

  Lock_Structure_Size = 10M * 4 * 8 = 320M

  Note: Minimum size of 13M is recommended

Performance and Tuning – Lock Structures

- Small lock structures result in increased false contention rates
  - True or false contention results in asynchronous lock requests
- To review false contention
  - RMF CF activity report
  - SMF 42 subtype 17
  - D SMS,CFLS command
- Recommended false contention rate is < 5%

Performance and Tuning – Data Class

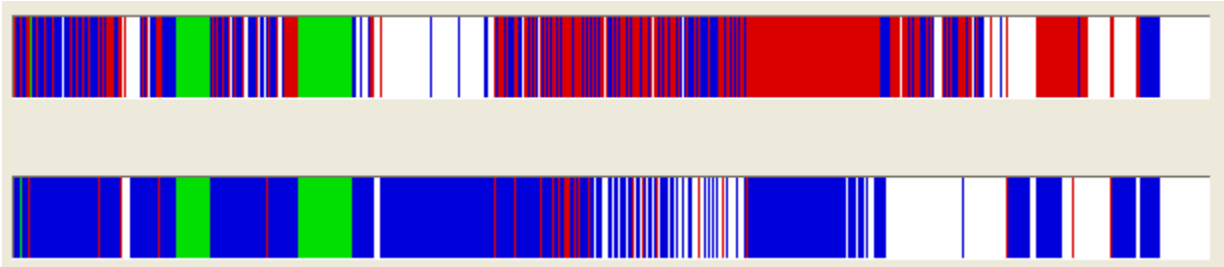
- BWO (NO | TYPECICS | TYPEIMS)
  - BWO: Backup While Open
  - NO: BWO not used for CICS or IMS VSAM data sets
  - TYPECICS: BWO used for CICS VSAM file control data sets
    - Requires multiple SVC 26 calls to catalog
    - Component 1 lock held across VRM requests
  - TYPEIMS: BWO used for IMS VSAM data sets with IMS 6.1+
  - Recommendation: To maximize concurrent VRM performance, only use BWO when absolutely necessary.
- RLS CF Cache Value (ALL | NONE | UPDATESONLY)
  - ALL: Cache data and index CIs
  - None: Cache index CIs only
  - Updates: Cache CIs for write requests only
  - This setting only applies with RlsMaxCFFeatureLevel(A)!

Performance and Tuning – Data Class

- RLS Above the 2-GB Bar (YES | NO)
  - Specifies whether SMSVSAM can use 64-bit virtual buffering for the data set
  - NO:
    - Recommended for heavy insert and update data sets
  - YES:
    - Recommended for heavy read data sets where data is reaccessed within one hour's time frame
  - Must specify a non-zero RLSAboveTheBarMaxPoolSize
  - The current design of 64 bit buffering uses 10-20% more CPU for equivalent 31 bit requests, however, the large 64 bit pool size allows for increased buffer hits over the 31 bit pool. 64 bit buffering will be redesigned in a future release to address this CPU increase.

Performance and Tuning – Data Class

- Extent Constraint Removal (YES | NO)
  - Specifies whether data sets can have > 255 extents
- CA Reclaim (YES | NO)
  - Specifies whether DASD space for empty CAs in KSDS data sets will be reused
  - Value saved in catalog
  - Only used if CA reclaim function enabled in IGDSMSxx or with SETSMS command
  - Recommended to avoid data set fragmentation and to reduce reorg frequency



Performance and Tuning – Data Set

- CISIZE
  - Currently, both RLS and base VSAM will serialize the *entire* data set when
    - CI / CA split is in progress
    - CI / CA reclaim is in progress
    - Accessing spanned records
  - Recommendation: When reasonable, increase CI size for data sets with higher insert activity.
    - Larger is better in terms of CPU
    - Larger also means larger buffers: performance vs. storage tradeoff
    - Larger CIs will contain more records which may lead to contention on DASD flush

Performance and Tuning – ACB / GENCB

- RLSREAD (NRI | CR | CRE)
  - NRI
    - No Read Integrity
    - Will not get record lock
  - CR
    - Consistent Read
    - Will get / release record lock
  - CRE
    - Consistent Read Extended
    - Will get record lock
    - Lock released at commit (recoverable data sets only)

Performance and Tuning – RPL

- OPTCD
  - ASY | SYN
    - Asynchronous: SRB
    - Synchronous: TCB
  - DIR | SEQ | SKP
    - Direct: Process request starting with specific record
    - Sequential: Process request record by record
    - Skip Sequential: Position on a record and move sequentially
  - NRI | CR | CRE
    - No Read Integrity
    - Consistent Read
    - Consistent Read Extended

Performance Measurements

Performance Measurements – SMF 64

- SMF 64 Records
  - Cut by CLOSE and EOV on a per ACB basis:
    - SMF64DLR: number of logical records
    - SMF64DDE: number of delete requests
    - SMF64DIN: number of insert requests
    - SMF64DUP: number of update requests
    - SMF64DRE: number of retrieve requests
    - SMF64BMH: number of BMF hits in the local buffer pool
    - SMF64CFH: number of CF hits in the RLS cache structure
    - SMF64RIO: number of requests read from DASD
    - SMF64DEP: total number of requests
    - SMF64NLR: number of logical records at open

Performance Measurements – SMF 42

- SMF 42 Subtypes 15, 16, 17, 18, 19
  - **Subytpe 15** - RLS statistics by storage class
  - **Subtype 16** - RLS statistics by data set
    - Must use V SMS,MONDS(spherename),ON to collect subtype 16 statistics
  - **Subtype 17** - RLS CF lock structure usage
  - **Subtype 18** - RLS CF caching statistics
  - **Subtype 19** - Buffer Manager LRU statistics
- Note: Only one system in the sysplex collects the SMF 42 records. The system collecting the records is displayed in the D SMS,SMSVSAM operator command.

Performance Measurements – RMF

- RMF Monitor III
  - All statistics gathered from interface call with SMSVSAM address spaces
  - Collected on RMF defined intervals
    - Can be matched to SMF intervals with SMF_TIME(YES)
  - Any SMSVSAM data in RMF monitor reports also found in SMF 42 records

Performance Measurements – RMF

- RLSSC & RLSDS
  - RLSSC: VSAM RLS activity by storage class
  - RLSDS: VSAM RLS activity by data set
  - Two different organizations split into sequential and direct requests
  - Provides information on:
    - Local buffer access (BMF)
    - CF cache structure hit %
    - DASD hit %
    - % valid buffer / cache hits
    - % false invalid buffer / cache hits
  - Assists in determining sizes
    - CF cache structures
    - Buffer pools
  - Helps analyze overall application performance

Performance Measurements – RMF

RMF Sysplex Report Selection Menu

Selection ===>

Enter selection number or command for desired report.

Sysplex Reports

1	SYSSUM	Sysplex performance summary	(SUM)
2	SYSRTD	Response time distribution	(RTD)
3	SYSWKM	Work Manager delays	(WKM)
4	SYSENQ	Sysplex-wide Enqueue delays	(ES)
10	RLSSC	VSAM RLS activity by storage class	(RLS)
11	RLSDS	VSAM RLS activity by data set	(RLD)
12	RLSLRU	VSAM LRU overview	(RLL)

RLSSC - Sysplex Totals View

```
                    RMF V1R8   VSAM RLS Activity  - SYSPLEX        Line 1 of 12
Command ===>                                                    Scroll == => HALF
Samples: 59      Systems: 2    Date: 10/31/06  Time: 13.16.00  Range: 60    sec
               < 2GB / > 2GB
LRU Status   :  Good / Accel
Contention % :   0.0 /  0.0
False Cont % :   0.0 /  0.0
Stor Class  Access  Resp   -------- Read ----------  ------ BMF -------  Write
                    Time    Rate  BMF%   CF%  DASD%  Valid%  False Inv%  Rate
RLS_SC1
 Below 2GB   DIR    0.004  665.6  88.2   0.5   11.3    100      0.01      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0      0.00      0.00
 Above 2GB   DIR    0.004  665.6  88.2   0.5   11.3    100      0.01      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0      0.00      0.00
RLS_SC2
 Below 2GB   DIR    0.005  200.0  90.5   0.0    9.5    100      0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0      0.00      0.00
```

RLSSC - System/CF Structure View

```
                    RMF V1R8 VSAM RLS Activity   - SYSPLEX        Line 1 of 23
Command ===>                                                  Scroll == => HALF
Samples: 120     Systems: 2    Date: 10/31/06  Time: 13.25.00  Range: 120   Sec
               < 2GB / > 2GB
LRU Status    :  Good / Accel  Storage Class : RLS_SC1
Contention % :    0.0 /  0.0   Cache Set     : PUBLIC1
False Cont % :    0.0 /  0.0
System/CF   Access  Resp    -------- Read ----------  ------ BMF -------  Write
                    Time     Rate  BMF%   CF%  DASD%  Valid%  False Inv%  Rate
*ALL
  Below 2GB  DIR    0.000  14.98  83.0   0.0   17.0    100       0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0       0.00      0.00
  Above 2GB  DIR    0.000  14.98  83.0   0.0   17.0    100       0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0       0.00      0.00
SYS1
 CACHE01
  Below 2GB  DIR    0.000   7.49  83.0   0.0   17.0    100       0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0       0.00      0.00
  Above 2GB  DIR    0.000   7.49  83.0   0.0   17.0    100       0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0       0.00      0.00
```

RLSDS - Sysplex Totals View

```
                    RMF V1R8 VSAM RLS Activity   - SYSPLEX        Line 1 of 20
Command ===>                                                    Scroll == => HALF
Samples: 120     Systems: 2    Date: 10/31/06  Time: 13.25.00  Range: 120  Sec
               < 2GB / > 2GB
LRU Status   :  Good / Accel
Contention % :   0.0 /  0.0
False Cont % :   0.0 /  0.0
Sphere/DS   Access  Resp   -------- Read ----------  ------ BMF -------  Write
                    Time    Rate  BMF%   CF%  DASD%  Valid%  False Inv%  Rate
BMAI.VSAMIN.MEGA
 BMAI.VSAMIN.MEGA.AIX.DATA
  Below 2GB  DIR    0.003   0.01   0.0   0.0    100    0.0       0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0       0.00      0.00
  Above 2GB  DIR    0.003   0.01   0.0   0.0    100    0.0       0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0       0.00      0.00
 BMAI.VSAMIN.MEGA.AIX.INDEX
  Below 2GB  DIR    0.003   0.03  50.0   0.0   50.0    100       0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0       0.00      0.00
  Above 2GB  DIR    0.003   0.03  50.0   0.0   50.0    100       0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0       0.00      0.00
```

RLSDS - System View

```
                 RMF V1R8 VSAM RLS Activity   - SYSPLEX         Line 1 of 9
Command ===>                                                  Scroll == => HALF
Samples: 120     Systems: 2    Date: 10/31/06  Time: 13.25.00  Range: 120  Sec
               < 2GB / > 2GB
LRU Status   : Good  / Accel                        Storage Class : RLS_SC1
Contention % :   0.0 /  0.0                 Cache Set     : RLSCSET
False Cont % :   0.0 /  0.0              Data Set      : RLSADSW.VF01D.INVENTOR.IND
System/CF   Access  Resp    -------- Read ----------  ------ BMF -------  Write
                    Time     Rate  BMF%   CF%  DASD%  Valid%  False Inv%  Rate
*ALL
 Below 2GB   DIR    0.000  24.27  97.7   2.2    0.1    100      0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0      0.00      0.00
 Above 2GB   DIR    0.000  24.27  97.7   2.2    0.1    100      0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0      0.00      0.00
SYS1
 RLSCACHE01
 Below 2GB   DIR    0.000   0.20   100   0.0    0.0    100      0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0      0.00      0.00
SYS2
 RLSCACHE01
 Above 2GB   DIR    0.000  24.07  97.7   2.2    0.1    100      0.00      0.00
             SEQ    0.000   0.00   0.0   0.0    0.0    0.0      0.00      0.00
```

Performance Measurements – RMF

- RLSLRU
  - Provides buffer manager LRU statistics for each system
  - Helps to size 32 and 64 bit pools
  - Properly sized pools help to tune cache structures

Performance Measurements – RMF

- RLSLRU

An example of a RLSMAXPOOLSIZE = 500 and RLSABOVETHEBARMAXPOOLSIZE = 2000 specification:

```
                    RMF V1R8 VSAM LRU Overview   - SYSPLEX          Line 1 of 2
Command ===>                                                   Scroll == => HALF
Samples: 120     Systems: 2    Date: 10/31/06  Time: 13.25.00  Range: 120   Sec
MVS          Avg CPU  - Buffer Size -  Accel  Reclaim  ------ Read -----
System        Time      Goal   High      %       %     BMF%   CF%  DASD%
SYS1
 Below 2GB  0.023        500     352    0.0     0.0    82.0   8.0   10.0
 Above 2GB  3.543        2000    1552   0.0     0.0    95.0   2.0    3.0
SYS2
 Below 2GB  4.457        500     612    100.0   0.0    31.2   10.5   58.3
```

RLSLRU - Buffer Counts by Pool View

```
                RMF VSAM LRU Overview - Buffer Counts by Pool
The following details are available for MVS System: SYS2
Press Enter to return to the Report panel.
Fixed Storage  Below 2 GB :     0    Above 2 GB :     0
Real Storage % Below 2 GB :     0    Above 2 GB :     0
----------------- Below 2 GB ----  ---- Above 2 GB ----
                 Low   High    Avg     Low   High    Avg
Fixed Pages        0      0      0       0      0      0
Buffer Counts by Pool:
                                                  More: +
  2K             163    318    226       0      0      0
  4K             713   1537   1299       0      0      0
  6K               0      0      0       0      0      0
  8K               0      0      0       0      0      0
 10K               0      0      0       0      0      0
 12K               0      0      0       0      0      0
 14K               0   5000      0       0      0      0
 16K             460    678    656       0      0      0
 . . .
 30K               0   9000      0       0      0      0
 32K               0   7949      0       0      0      0
```

#SHAREorg

SHARE
Technology • Connections • Results

SMSVSAM Diagnostics

SMSVSAM Diag – Display Commands

- D SMS,CFLS(lock_structure)
  - Displays lock structure statistics to the console
  - Contention values for IGWLOCK00 (default) or specified structure
  - Tracks
    - LockRate: Number of lock requests / sec
    - ContRate: Percentage of requests that encountered contention
    - FContRate: Percentage of requests encountering false contention
    - WaitQLen: Average number of requests waiting behind locks

SMSVSAM Diag – Display Commands

- D SMS,CFLS(lock_structure)

```
IGW320I 00:01:48 Display SMS,CFLS(IGWLOCK00 )
PRIMARY STRUCTURE:IGWLOCK00 VERSION:C57C859902B9264E SIZE:95232K
RECORD TABLE ENTRIES:34438 USED:747
SECONDARY STRUCTURE:IGWLOCK00 VERSION:C58548B866A7576E SIZE:95232K
RECORD TABLE ENTRIES:34438 USED:747
LOCK STRUCTURE MODE: DUPLEXED  STATUS: ENABLE
System   Interval   LockRate   ContRate  FContRate   WaitQLen
J90      1 Minute       15.1      0.000      0.331       0.00
J90       1 Hour        11.9      0.000      0.047       0.01
J90       8 Hour         8.2      0.001      0.039       0.00
J90        1 Day        17.3      0.001      0.042       0.01
  (09)   1 Minute        7.5      0.000      0.161       0.00
  (09)    1 Hour         7.0      0.000      0.020       0.00
  (09)    8 Hour         6.6      0.001      0.033       0.00
  (09)     1 Day        20.9      0.001      0.029       0.00

***************** LEGEND ******************
  LockRate = number of lock requests per second
  CONTRATE = % of lock requests globally managed
  FCONTRATE = % of lock requests falsely globally managed
  WaitQLen = Average number of requests waiting for locks
```

SMSVSAM Diag – Display Commands

- D SMS,SMSVSAM,DIAG(Contention)
  - Displays latch contention on SMSVSAM resources
  - Latches are 8 bytes of storage used to serialize system resources
    - Two states: Holder & Waiter
    - Latches are not shared
  - The contention display shows
    - Latch address
    - Holder and any waiters if the latch is in contention
    - Elapsed time of contention period
  - System specific (will not show latches around the plex)
  - Elapsed time of more than a few seconds is likely a problem
  - Should be used in conjunction with D GRS,C

SMSVSAM Diag – Display Commands

- D SMS,SMSVSAM,DIAG(Contention)

```
 09.55.29 SYSTEM1           IGW343I VSAM RLS DIAG STATUS (V.01)
|---RESOURCE----| |------ WAITER ------| |--HOLDER---| ELAPSED
 TYPE        ID    JOB NAME ASID   TASK   ASID   TASK     TIME
 -------- -------- -------- ---- -------- ---- -------- --------
 LATCH    7F158C70 SMSVSAM  003A 008DA250 003A 008D7218 00:00:06
     DESCRIPTION: IGWLYSPH - SHM OBJECT POOL
 LATCH    7F151E78 SMSVSAM  003A 008D7218 003A 008DC1C8 00:00:21
     DESCRIPTION: IGWLYDTS - SHM OBJECT POOL
 LATCH    7BAD43B8 SMSVSAM  003A 008DC1C8 002D 007F3000 00:19:09
 LATCH    7BAD43B8 SMSVSAM  003A 008D5A48 002D 007F3000 00:22:09
 LATCH    7BAD43B8 SMSVSAM  003A 008D6938 002D 007F3000 00:33:23
 LATCH    07F1B1D0 SMSVSAM  003A 008D64F8 003A 008D6CF0 01:47:20
 LATCH    07F1D3B8 SMSVSAM  003A 008D6CF0 0000 00000000 11:23:30
```

SMSVSAM Diag – Display Commands

- D SMS,SMSVSAM,QUIESCE
  - Displays outstanding quiesce activity for SMSVSAM
  - A Quiesce event is a decision to shutdown / allow CICS access to an RLS dataset
  - Any CICS regions participating in the event will be in the display
  - If there is not an event in progress, IGW540I rejects the command
  - Invaluable to determine CICS regions holding up a quiesce request
  - CICS regions with 00.00.00 in the “completed” section are likely in trouble
    - Dump these regions with SMSVSAM prior to termination
  - Often the region can be terminated instead of SMSVSAM
    - Allows the quiesce event to finish
    - Saves a plex-wide SMSVSAM outage

SMSVSAM Diag – Display Commands

- D SMS,SMSVSAM,QUIESCE
  - With quiesce activity

```
IGW540I 13.30.45 DISPLAY SMS,SMSVSAM,QUIESCE
  MVS1
   SPHERE NAME: DLLEHR.TEST1
   SYSTEM NAME: MVS1     START TIME:   .27.50 TOTAL ELAPSE TIME: 57.02.55
   PARTICIPATING SUB-SYSTEM STATUS:      SCHEDULED:  COMPLETED:  ELAPSE:
    SUB-SYSTEM NAME:  CI1AORP1            .27.50     00.00.00   57.02.55
    SUB-SYSTEM NAME:  CI1AORP2            .27.50     .27.50     00.00.00
```

  - Without quiesce activity

```
IGW540I 07.54.28 DISPLAY SMS,SMSVSAM,QUIESCE
DISPLAY SMSVSAM QUIESCE SPHERE IS REJECTED.
NO QUIESCE EVENTS ARE ACTIVE ON THIS SYSTEM.
```

SMSVSAM Diag – Hang / Wait in RLS

- Hang / wait conditions in RLS
  - May be related to activities such as
    - Accessing data sets for RLS
    - Starting / stopping of RLS client spaces (e.g. CICS)
    - Starting / stopping of SMSVSAM
    - SMSVSAM commands not responding
  - Detecting / pinpointing this condition requires
    - Display command output
    - Plex-wide dumps of SMSVSAM and related client spaces

SMSVSAM Diag – Hang / Wait in RLS

- Display commands to run
  - D SMS,SMSVSAM,ALL
  - D SMS,TRANVSAM,ALL (if TVS is in use)
  - D SMS,SMSVSAM,DIAG(CONTENTION) (system scope)
  - D SMS,SMSVSAM,QUIESCE (system scope)
  - D GRS,C (system scope)
  - D XCF,STR,STRNM=[IGWLOCK00 | lock_structure]
  - D XCF,STR,STRNM=[cachestructurename]

SMSVSAM Diag – Hang / Wait in RLS

- Plex-wide dumps
  - Most required info is in SMSVSAM address and data spaces
    - Both SMSVSAM and MMFSTUFF dataspaces are needed
  - SMSVSAM is thoroughly plex-wide so the plex must be dumped
    - Creating an IEADMCxx parmlib member is recommended
    - Example of a plex-wide parmlib definition:

```
COMM=(SMSVSAM hang)
JOBNAME=(*MASTER*,SMSVSAM),
DSPNAME=('SMSVSAM'.*),
SDATA=(ALLNUC,COUPLE,CSA,GRSQ,LPA,LSQA,PSA,RGN,SUM,SQA,SWA,TRT,XESDATA),
REMOTE=(SYSLIST=(*('SMSVSAM')),DSPNAME,SDATA)
```

SMSVSAM Diag – Internal Logic Errors

- Internal logic errors
  - Surfaced as an ABEND0F4 with unique RC and RSN
  - SVC dump taken (typically plex-wide)
  - Depending on severity, processing may be stopped
  - Dumps may be accompanied by syslog flower box
    - IGW400I / IGW405I
    - VRM request type and key
    - Jobname / data set name
    - ASID / TCB
    - User and data space RPL
    - VRM footsteps
  - When encountered, necessary doc includes
    - Plex-wide 0F4 dumps
    - OPERLOG
    - Plex-wide software detail EREP reports (LOGREC)

RLS Maintenance and References

Maintenance and References – PSP Bucket

- The RLS PSP bucket is on the web at:
  - http://www14.software.ibm.com/webapp/set2/psearch/search?domain=psp
- 'Search for: SMSVSAM'
- Maintenance organized by release
  - Upgrade ZOSV1R11, Subset DFSMS
  - Upgrade ZOSV1R12, Subset DFSMS
  - Upgrade ZOSV1R13, Subset DFSMS

References

- VSAM Demystified
  - SG24-6105-01
- z/OS V1R13 DFSMSdfp Storage Administration
  - SC26-7402-16
- z/OS V1R13.0 DFSMS: Using Data Sets
  - SC26-7410-11
- z/OS MVS System Management Facilities
  - SA22-7630-24

Acknowledgments

- Special Thanks to the following people for their help
  - DFSMS RLS Level 2 and Development teams
  - Doug Lehr for much of the basis of this presentation
  - Terri Menendez for her work on the Buffer Performance and diagnostic guidelines
  - Kathleen Ting and Pooja Bhargava for their work with RMF III
  - Our awesome clients who trust us with their data and help to ensure value is regularly driven back into the product!

Conclusion

- RLS performance increases as a result of shortening the I/O path.
- Reducing time in I/O can be best accomplished when
  - Data is in the buffers
  - Data is in the cache structures on the coupling facility
- Reducing CPU utilization at the data set level is another important means of performance improvement.
  - Better data set tuning
  - Avoid serialization intensive situations
  - Take advantage of offerings in the SMS configuration

#SHAREorg

SHARE
Technology · Connections · Results

DFSMS Basics: VSAM

VSAM RLS Best Practices

How to leverage VSAM RLS best practices and performance in your environment

David LeGendre, dlegendr@us.ibm.com

Session : 10967