Hints and Tips for Improving your DFSMShsm Environment

Chris Taylor
IBM Corporation

Steven Clar
Rocket Software

Tuesday, August 3, 2010
Session Number 8026

SHARE in Boston

Trademarks

The following are trademarks of the *International Business Machines Corporation:*

IBM, DFSMS/MVS, DFSMShsm, DFSMSrmm, DFSMSdss, DFSMSopt, DFSMS Optimizer, z/OS, eServer, zSeries, MVS, FlashCopy®

The information contained in this presentation is distributed on an 'AS IS' basis without any warranty
either expressed or implied, including, but not limited to, the implied warranties of merchantability or fitness for a particular purpose.   The use of this information is a customer responsibility and depends on the customer's ability to evaluate and integrate it into the customer's operational environment.

SHARE in Boston

Agenda Topics

- ## Setting the stage
  - ### How these Tuning Tips came to be
    - IBM's DFSMShsm Health Assessment
  - ### What tuning tips will be covered
    - HSM Log Files
    - Recycle
    - Expirebv
    - HSM Patches

- Fast Subsequent Migration
- When to split CDS's
- Enabling RLS
- Migration/Backup Errors
- Compaction
- Interval Migration
- Cancelling DFSMShsm
- Home Grown Reporting
- Tuning CDS's
- Reorganizing Control Data Sets

Setting the Stage

- The discussion you are about to be part of is real. The information being presented was gathered while performing IBM DFSMShsm Health Assessments at different customer sites.

- These sites ranged from small to large shops with a couple hundred thousand to millions of DFSMShsm managed migration and backup records and tapes.

- In most assessments performed sites or administrators are not aware critical SETSYS values are disabled or not active.

Setting the Stage (continued)

- In most instances SETSYS values and SMS/HSM configurations were performed years ago and usually never revisited or modified over the years.
    - Original administrators are no longer there and environment has been inherited
    - "Not sure why, but we dare not change it"

- These scenarios could be costing companies a lot of extra HSM resources and dollars.

- These are not errors, rather examples showing why it is important to have a toolset that can show you where change may be needed and why.

HSM Log files

- HSM Logfiles are used to track HSM activity
- Required if using an ISV solution that reads the logs
- Disable if not needed
  - Additional overhead when using logging
- One set of log files per HSM started task
- Active logfile is always the HSMLOGX dataset
  - Exclusive enqueue issued by HSM started task
  - HSM swaps the log files by renaming them
    - *Need to reside on the same disk volume*
- HSMLOGY data set can be analyzed using ARCPRLOG/ARCPEDIT programs

HSM Log files (cont.)

- HSMY logfile has EOF marker written after close
  - No EOF marker when being used as active HSMLOGX

- Some products have the ability to read past current data
  - Some data may therefore be old

- Some customers process log files and then "zero" them
  - Raw log data is not available for further analysis

- Recommendation: Write the HSMLOGY file to a GDG to allow further analysis before "zeroing"

Log Files (cont.)

- Check DISP in HSM started procedure
  - We often see DISP=SHR or OLD coded

- If LOGSW=YES, then DISP must be OLD
  - PDA data sets are swapped automatically at startup

- If LOGSW=NO, then DISP should be MOD
  - If DISP=OLD or SHR, data will be overwritten at HSM restart

Recycle

- Recycle purpose is to free-up HSM tape by merging no longer needed migrated and backup data to new tape releasing the cycled tape back to the tape management system for reuse.
  - Also used to move HSM data to new tape devices

- Truth or Myth –
  - Virtual Tape Library's
    - *Some sites are under the impression because they have HSM tapes in a VTL they have unlimited tapes available.*
      - *Myth – as the saying goes, you are only as strong as the weakest link. Though there could thousands of tapes defined, the weak link is the back end storage.*

Recycle

- Truth or Myth (continued) –
  - We aggressively migrate HSM data to the VTL because it uses less resources than physical tape.
    - Truth and Myth –
      - *VTL's do process data faster than physical tape, aside from the regular DFSMShsm resources for recycle a VTL also uses these additional cycles –*
      - *Cache for loading data*
      - *Loading of complete tape into memory*
      - *Hardware decompression*
    - Recommended recycle percentage
      - *Physical Tape – 30-40 percent valid data*
      - *Virtual Tape – 10-20 percent valid data*

EXPIREBV – Scenario 1

- Number of backup data set versions was more than 10 times the number of migrated data sets
  - Customer was expecting to keep maybe 2 backup versions

- Result

  - Expirebv had never been run at this site

  - Ran Expirebv in DISPLAY mode

  - Estimated that instead of 78,000 backup versions, 8,000 would remain after 30 days

EXPIREBV – Scenario 2

- Customer complained that one BCDS component needed to be reorged more often than others
- Ensured that EXPIREBV was running every day

- Result
  - Customer was keeping all backup versions for 7 years after migration version was deleted
  - By reducing the retention of test data to 1 year could save
    - 28 % of backup space occupied on tape (1.53 TB)
    - 40% of number of BCDS entries

EXPIREBV

Purpose -

- Delete expired DFSMShsm backup versions of datasets, based on Management Class or supplied values
  - Management Class – Retain Days extra backup version
  - Also deletes expired ABARS versions
- Flags when a data set is _detected_ as uncataloged
  - Either on primary disk or migrated version
- Updates BCDS to store date of Expirebv run
- Subsequent Expirebv run deletes backup version
  - Based on Management Class – Retain Days only backup version

Patches

- Patches are applied via the DFSMShsm PATCH command
  - Often included in the ARCCMDxx member

- Supplies functionality not available using standard HSM commands

- Some patches are for tuning, others are for diagnostics

- Supported Patches are documented in the DFSMShsm Implementation and Customization Guide

Patches - Tuning

- Examples
  - PATCH .MCVT.+52 bits(..1.....)
    - Alter the default for tape data set recalls from a wait state to a no-wait state
    - Allows TSO users to perform other tasks while recall takes place in the background

  - PATCH .MCVT.+321 'SHORT ' VERIFY(.MCVT.+321 'HSMACT')
    - Change HLQ of HSM DASD activity log from HSMACT to SHORT

Patches – Problem Determination

- Examples
  - PATCH .MGCB.+26 X'FF'
    - Used to determine why an SMS-managed data set is not selected during volume migration
  - PATCH .BGCB.+24 X'FF'
    - Used to determine why SMS-managed data sets are not being selected during volume backup
- These patches produce a <u>lot</u> of messages
  - ARC1245I with Reason Codes GT 90 for migrations
  - ARC1334I with Reason Codes GT 90 for backups

- We discovered this because the client's error numbers were very high
- Use diagnostic patches only when needed or directed by Level 2 support

Fast Subsequent Migration

- Discussed in Best Practices Session
- We found very few sites using it
- Easy to implement
  - SETSYS TAPEMIGRATION(RECONNECT(**NONE** | ALL | ML2DIRECT))
- Reduces need for data movement
- Review SETSYS MIGRATIONCLEANUPDAYS
  - E.g. SETSYS MIGRATIONCLEANUPDAYS(30 30 10)
    - 1st parm: Number of days to keep MCDS record after recall
    - 2nd parm: Number of days to keep daily & volume statistic recs.
    - 3rd parm: Number of days to retain records for reconnection

Thrashing

- Thrashing can be described in 2 ways
  - A data set which is migrated and recalled within a few days
  - Data sets which are migrated and recalled multiple times
- Often generation data sets involved
  - Management Class says to allows GDS early migration
    - MC Class field # GDG Elements on Primary
  - Some jobs recall entire GDG rather than relative generation
    - Data is recalled even when not needed
- Consider not migrating small datasets
  - Migration may not be worth the processing overhead
  - Use ARCMDEXT exit to exclude from migration
    - Can also allow migration to ML1 but exclude from ML2

Thrashing

- HSM SMF records (FSR) can be used to look for thrashing
  - Products can also be used

```
DFHSM Recall Thrashing Report - Detailed        07/16/10        16:36:17

2010/07/16  : Daily Report

Sid        Time    Data Set Name                                        MigDays   Mclas
----    --------   ------------------------------------------------    -------   -----
SYS1    13:01:00   P390.MIGR.AUDIT.D071510                                   0   MCDEF
SYS1    15:39:01   BJT230.GLOBAL.CSI                                         4   MCPRD
SYS1    15:39:12   BJT230.GLOBAL.SMPPTS                                      4   MCPRD
SYS1    15:44:37   GLO310.GLOBAL.CSI                                         4   MCPRD
SYS1    15:44:41   HOSM612.GLOBAL.CSI                                        0   MCPRD
SYS1    15:44:50   ITM621.GLOBAL.CSI                                         0   MCPRD
SYS1    15:45:11   VTM300.GLOBAL.CSI                                         0   MCPRD
SYS1    15:53:37   SYS2.BJT230.SMPE.CNTL                                     4   MCDEF
SYS1    16:07:59   GLO310.SMPPTS                                             0   MCPRD
SYS1    16:11:52   GLO310.SMPLOG                                             0   MCPRD
SYS1    16:11:54   GLO310.SMPMTS                                             0   MCPRD
SYS1    16:20:24   VTM300.SMPMTS                                             0   MCPRD
```

- Note: FSR records can also include data sets processed for extent reduction
  - These are not really trashing
  - Review SETSYS MAXEXTENTS

Thrashing – IEFBR14

- Production jobs often use IEFBR14 with DISP=(x,DELETE) as first step
- HSM will recall the data set in order to delete
- z/OS V1R11 allows data sets to be deleted without Recall
- Changes in ALLOCxx member in SYS1.PARMLIB
  - SYSTEM IEFBR14_DELMIGDS(NORECALL)
  - Default value is LEGACY
- Recommend NORECALL unless another product already being used
  - e.g. zOSEM

Guaranteed Backup Frequency

- Storage Group parameter
- Specifies the maximum number of days between backups
  - Even if it was not modified
- New backup copy replaces and invalidates old copy
  - Additional I/O required
    - Verify last backup date in BCDS
    - Backup data
    - Create new version in BCDS
    - Delete old version from BCDS
- More recycles required with larger capacity tapes
- Recommend using this feature sparingly

When to Split a CDS

Purpose -

- ## Increase xCDS size
    - More data and records can be managed by HSM.
    - Increases HSM performance.

Things to know -

- ## Provides Multi-Cluster CDS
    - Increases HSM performance going across multiple splits than one cluster.
    - Manage more HSM owned data in MCDS and BCDS

VSAM Record Level Sharing for DFSMShsm CDSs

- VSAM RLS can greatly improve throughput and decrease windows
  - Also improves Audit performance
- Requires Coupling Facility access from DFSMShsm hosts
- If one DFSMShsm host accesses in RLS mode, <u>all</u> hosts must access the CDSs in the same mode
  - HSM startup will fail if an incompatible serialization mode is found
- Control Data Sets must be SMS-managed
- If already using VSAM RLS for other applications, use pre-existing structures

VSAM Record Level Sharing Enablement

- Steps for setting up VSAM RLS (Very high-level)
  - Define new CFRM policy in coupling facility
  - Define 3 Share Control data sets (2 active, 1 spare)
  - Add Shared CDSs to SMSVSAM
  - ISMF
    - Define Cache sets
    - Storage class changes
    - Data class changes
  - Changes to IGDSMSxx
    - E.G. RLSINIT(YES)

Using RLS with DFSMShsm

- Steps to be taken to enable HSM to open the CDSs in RLS mode
  - Control data sets must be altered to change LOG(NONE)
  - DFSMShsm startup proc change to CDSSHR=RLS

- DFSMSdss must be used for CDS backups
  - CDSVERSIONBACKUP DATAMOVER(DSS)

- To fallback
  - Stop HSM tasks on all LPARs
  - ALTER cdsname NULLIFY(LOG)
  - CDSSHR=YES

- More detailed info in Session 3045 from Share in San Jose
  - VSAM RLS Best Practices session 8062 on Wednesday at 130pm

Common Recall Queue - CRQ

- Discussed in Best Practices Session
- Consolidates recall requests and spreads them across HSM instances
  - Balances workloads around the complete HSMplex
- Allows important recalls to be prioritized ahead of lesser ones
  - DFSMShsm exit ARCRPEXT
- Optimizes Tape mounts
  - Single tape mount satisfies requests from multiple LPARS
- Requests can be carried out by all or some of the systems
  - Allows systems without attached tape to issue recall requests

Common Recall Queue – Enablement

- Requires Coupling Facility
- I & C guides recommends initial values
  - INITSIZE= 5120KB
  - SIZE=10240KB
    - Allows between 3900 and 8400 <u>concurrent</u> recalls
    - Recommendation assumes that 33% request a unique ML2 tape
    - Also use http://www-947.ibm.com/systems/support/z/cfsizer/crq/
- Use IXCMIAPU to define the structure to the CFRM
- Activate new structure
  - SETXCF START,ALTER,STRNAME=name,SIZE=xxxx
- Issue HSM command
  - SETSYS COMMONQUEUE (RECALL(CONNECT(basename)))

Common Causes of Migration & Backup Failures

- Everyday in most shops DFSMShsm primary, secondary and backup are run at specific times daily. In most cases business's have grown, storage farms have grown and managed data has grown, but when was the last time your scheduled tasks were reviewed or verified?

- Here are some common failures documented while doing the Health Assessments.

  - Data Set in Use (migration/backup) –
    - A common encountered error, everyday DFSMShsm will try to migrate and backup these data sets and fail.
      - *Waste of DFSMShsm resources*

Common Causes of Migration & Backup Failures

- Common causes of migration/backup failures (continued) –
  - No space on ML1 Volume
    - This is usually seen with large data sets. Some simple solutions include adding additional volumes to ML1 pool, modifying management rules to expire more data on primary pools, setting up automation rules to move failed migration directly to ML2.
      - *Waste of DFSMShsm resources*

  - HSM Backup Critical Errors (condition code ne 0) –
    - HSM backup is critical to shops using this as their first level data recovery.
      - *Backup window overlaps batch processing*
      - *Ctlg errors (rc30) / DFDSS errors (rc68) / vtoc discrepancy (rc87)*
        - *Waste of DFSMShsm resources*

Common Causes of Migration & Backup Failures

- Common causes of migration/backup failures (continued) –
  - The most common migration and backup error noted when doing the Health Assessments were for Unsupported Datasets (rc99, rsn04).
    - Cause of the problem is incorrectly defined data sets (no DSORG).
    - Everyday HSM will try to migrate/backup these data sets and fail. We have seen situations where the same data sets have been failing for nine years and more. The quickest and easiest correction is to update the SMS routines to automatically assign a data class.
      - *Waste of DFSMShsm resources.*

Compaction – Scenario

- Customer reported that DFSMShsm was one of the 2 largest CPU users when automatic functions were running

  - Sometimes had to hold the running function in order to allow other work to proceed

- Recommendation was made to only use compaction for ML1 or backups to disk volumes, not tape

  - SETSYS COMPACT(DASDMIGRATE,DASDBACKUP)
  - SETSYS COMPACT(NOTAPEMIGRATE,NOTAPEBACKUP)

Compaction

- DFSMShsm provides software compaction using the Huffman algorithm

- Compaction can be specified for migration, backup, DASD and/or tape data sets

- Compaction is a CPU-intensive operation

- Ensure that compaction is not enabled for tape data sets
    - Use SETSYS TAPEHARDWARECOMPACT instead

Migration and SMS Storage Group Thresholds

- Found sites using unrealistic storage group thresholds
  - E.g. High threshold 80%, low threshold 1%
- Primary Space Management will attempt to process down to low threshold
- Interval Migration starts after halfway between high- & low-threshold is exceeded
  - Ends at low-threshold
- Leads to excessive cycles and missed space management windows
- Set values that are realistic for the storage group

Stopping a DFSMShsm Started Task

- Normal Shutdown
    - Operator issues shutdown of the DFSMShsm started task
        - HSM clears and updates ghost records.
        - HSM finishes and marks open tapes as completed.
        - HSM completes all remaining requests and scheduled work.
        - HSM finalizes all chained records.

Stopping a DFSMShsm Started Task

- Cancel HSM task
  - Ghost records are lost.
  - Tapes are left in FAILED status.
  - Chained records are left orphaned and disconnected.
  - Data at HIGH risk for unavailability.

Homegrown Reporting Tools

- Purpose –
    - Report on DFSMShsm activity
        - Migration, backup, recall, recover, extent reduction, PSM, SSM, etc.
            - *What and Why?*
            - *Age, times, etc.*
    - Successful/unsuccessful
        - What, Why and How long?
- Things to know –
    - Requires in-depth knowledge of HSM and records.
    - Requires other OEM software license
        - SAS
        - What is the plan for support and knowledge transfer?
- DFSMS Report Generator
    - See Session 2466 from Share in Seattle

Tuning of DFSMShsm

- In many shops DFSMShsm parameters have never been reviewed, analyzed or modified since first activated
  - Business's have grown, storage farms have grown, managed data has grown
    - SMS and HSM routines are usually added to but never reassessed
- What should you do -
  - Understand the work DFSMShsm is doing
    - What is it costing for scheduled tasks?
      - *Successful/Unsuccessful*
    - Are users wasting HSM resources for unnecessary work?
      - *What is the cost?*

Tuning of DFSMShsm

- Tuning DFSMShsm (continued) –
  - Tuning of DFSMShsm is always overlooked but should be done every so many years or when there is an addition of data (merging of new HSM data)

  - When to review SMS and HSM criteria –
    - When Companies grow.
      - *A review should be scheduled at minimum annually or semi-annually*
    - Businesses bought and sold
    - HSM managed data has grown exponentially

Tuning of DFSMShsm

- Tuning DFSMShsm (continued) –
  - An extremely helpful solution is to look for a product that will assist you with knowing HSM costs for better tuning

HSM Action	Return Code	Records Count	Tracks Read	CPU .01 Seconds	Recycle # Tapes	Estimated Cost
BACK-UP	0	930	366083	8211		20.53
BACK-UP	19	152	0	13		.03
EXBACKV	0	1013	0	130		.33
EXPIRED	70	36	0	0		.00
MIGRATE	0	123	12350	762		1.91
MIGRATE	6	190	0	11		.03
MIGRATE	24	1	0	0		.00
MIGRATE	37	118	0	7		.02
MIGRATE	58	79	0	4		.01
PARTREL	0	57	2657	12		.03
RECALL	0	95	4357	798		2.00
RECYCLE	0	9586	80520	1298	15	40.72

Reorganizing Control Data Sets

- Should you reorganize a Control Data Set?
  - Some Say Yes, Some Say No

- Why are you Reorganizing a Control Data Set?
  - Receiving warning messages from DFSMShsm
    - Incorrect sizing
    - Single cluster at 4Gb limit
      - *If using standard format VSAM*
    - "That is what we have been doing for years"

Reorganizing Control Data Sets

- Should you reorganize a Control Data Set (cont) ?

- Think of DFSMShsm as a crucial part of the OS
    - Every minute it is down –
        - Migrated data cannot be recalled
            - *Production delays*
        - Backed up data cannot be recovered
            - *Simple HSM lost revenue calculation*
                - *Total seconds HSM is working times cpu second divided by time HSM is down equals lost HSM revenue*

Reorganizing Control Data Sets

- Should you reorganize a Control Data Set (cont) ?
  - Look for alternative solutions
    - Correct sizing of CDSs
    - Reorg While Active products
    - Review why and if needed use tested procedure
  - Is there a performance increase after a reorg?
    - Yes, the reorg removes all splits, but when HSM is restarted the first thing it will do inside a CDS is a split.
    - Performance impact for a number of weeks
  - z/OS 1.12 New Feature VSAM CA Reclaim
    - Session 8054 discusses this new feature

Summary

- HSM Log Files
- Recycle
- Expirebv
- HSM Patches
- Fast Subsequent Migration
- When to split CDS's
- Enabling RLS
- Migration/Backup Errors
- Compaction
- Interval Migration
- Cancelling DFSMShsm
- Home Grown Reporting
- Tuning CDS's
- Reorganizing Control Data Sets