

The Care and Feeding of Your HSM Environment

OVERVIEW	3
HSM ACTIVITIES	3
Space management	4
Backup	4
Recycle	4
Dump	4
HSM ISSUES	5
High CPU usage	5
MAINTENANCE TASKS	6
Initial setups	6
Daily activities	8
HSM audit	9
MAINVIEW SRM REPORTING	9
Historical performance and capacity information	11
Application definitions	11
Space availability	11
Output management	12
Backup control	12
Recall allocation	12
SUMMARY	12
Helping you maintain advantage	13

OVERVIEW

Storage administrators are vital to IT departments because they protect and maintain a company's most valuable asset – its data. Today, storage administrators manage increasingly larger amounts of data. The combination of lower prices for storage devices and increasing government compliance regulations leads to more data being stored for longer periods of time.

One of the critical roles of the storage administrator is understanding and managing DFSMSHsm (better known as HSM). HSM provides backup, recovery, migration, and space management functions with optimum automation capabilities. HSM, which is shipped with z/OS, provides the following features:

- > ensures that space is available on DASD volumes, so that users can extend data sets and allocate new ones
- > ensures that backup copies of data sets are available in case working copies are lost or corrupted
- > automates manual storage management tasks
- > improves DASD usage by managing space and data availability in a storage hierarchy
- > provides a way to manage a data set throughout its lifecycle—from the time it's created until its last backup is no longer needed

Storage administrators can set and apply rules in HSM to manage storage environments. HSM enables you to manage storage at the data set level and device pool level.

This document outlines some best practices for working with HSM.

HSM ACTIVITIES

HSM performs the following activities:

- > space management (primary and secondary)
- > backup
- > recycle
- > dump

Space management

Space management ensures that only active data resides on DASD. With increasing pressures to reduce costs, it is important to manage space efficiently and migrate unnecessary data off of the primary DASD. To save space (and costs) the space management function automatically:

- > expires data sets that have exceeded expiration dates
- > moves unused data sets to where they can be stored less expensively or compresses them so that they require less space
- > frees unused portions of a data set (releases space)

Migration is the process of moving data from more expensive volumes to less expensive storage devices. When data is accessible to all users, it is considered level 0. Unreferenced level 0 data is migrated to a migration level 1 (ML1) disk, which is on DASD. This process is known as primary space management. The ML1 data is in a compressed format. Additional processes, such as data set compression, extent reduction, and space release, can occur during primary space management.

Secondary space management is the process of moving data from the ML1 disk to migration level 2 (ML2) tape. While ML2 volumes can reside on DASD, HSM supports only ML2 tape. Data on tape can expire over time, depending on the expiration date of the data set or the generation of more recent backup data sets.

Migrated data is *recalled* when a user needs it again. HSM generally processes recalls automatically.

Backup

The backup function backs up the HSM control data sets (CDSs), the journal, and all data sets whose change bit indicator has been flagged. These backups are generally used for local data set recovery. This data-set-level function determines whether to copy a data set, based on the backup frequency attribute of the storage group.

Recycle

Recycle is the process of moving valid backup or migration tape data to new tapes. As HSM expires the data residing on tape, the percentage of valid data that is residing on a specific tape is decreased. The process is to ensure that HSM is always maximizing its tape usage by using full, or nearly full, tapes.

Dump

The dump process dumps the entire contents of a disk volume to tape. Dumps are usually run weekly instead of daily and are normally sent offsite for business recovery purposes.

HSM ISSUES

Because HSM is so powerful and performs so many tasks, it has the potential to have problems. The areas of HSM that lead to problems are failures in space management, high CPU usage because of unneeded or wasteful actions, failures in the recycle process, internal control data set errors, and problems with the aggregate backup and recovery support (ABARS) and/or automatic dump processes.

Storage resource management (SRM) tools can help storage administrators examine and correct these problems:

- > high CPU usage
- > errors in control data sets
- > recycle failed
- > ABARS/Auto Dump failed

High CPU usage

High CPU usage generally raises red flags in IT shops. No one likes CPU resource hogs. By using an SRM tool, you can correct some of the most common causes of high CPU usage:

- > migrating unnecessary data sets
- > errors in space management
- > wasteful tape recycle
- > recall thrashing

Migrating unnecessary data sets. CPU usage can be attributed to many factors. An easy way to reduce CPU usage is to reduce or eliminate unnecessary HSM activity that is caused by ineffective management class policies or application JCL. SRM tools can help you reduce or eliminate the unnecessary HSM activity by tracking activity and thrashing and by tying the data sets to the DFSMS constructs.

Errors in space management. You can correct errors in space management by changing management class policies, removing undefined DSORG data sets, removing uncataloged data sets, and halting any other unnecessary HSM recall/migration/backup activity. For example, resolving return code 99, 19, 82 and 37 will give you quite a bit of productivity. Return code 99 is caused by an undefined DSORG, which can cause errors during backup and migration. Return code 37 is issued when there is not enough contiguous space to migrate the data set. If you receive a return code 37, either you can change the management class to prevent the large data set from going to ML1, or you can increase the amount of space available in the ML1 pool.

SRM tools enable you to quickly see which return codes you received and how to correct them.

Wasteful tape recycle. Wasteful tape recycle is usually a result of the *percent full* set too low. How the recycle is run—and when it is run—can also lead to wasteful recycles. Recycle should be part of an automated solution. It is best to run it during low tape drive usage, such as early in the morning just before normal business hours, or just before the nightly batch cycle runs.

Recall thrashing. Recall thrashing is a tremendous contributor to CPU cycles. Migrating and then immediately (or in a short period of time) recalling the same set of data sets wastes valuable HSM resources and can be quite costly. Understanding what is being migrated and quickly recalled requires a thrashing report. Examining the thrashing report and who or what is recalling the data sets will help you set better management class policies. You can also code batch JCL to be more efficient in how it handles GDG bases or sequential files.

Small data sets are the nemesis of HSM. Even when using small data set packing (SDSP) files, the constant migration of small data sets to ML1 and then to ML2 tape—only to be expired in a short period of time—is useless. The compaction of the files when migrating is minimal when compared to the cost of DASD. For SDSP data sets, take a closer look at when the data migrates and when it expires. Changing the management class policies for these data sets to let them “live and die” on primary disk is a good idea. It reduces unnecessary migrations and the subsequent expiration from HSM. Preventing the migration reduces the need to recycle tape volumes that have quickly expiring data.

MAINTENANCE TASKS

Maintaining HSM for optimum efficiency is essential. This section describes some of the tasks you need to complete to set up and maintain HSM.

- > initial setups
- > daily activity
- > HSM audit

Initial setups

Initial setups include creating the HSM PARMLIB, the SMS ACS routines, and the SMS constructs.

Space management cycles. To ensure that space management is running efficiently, determine when the primary and secondary space management cycles start and end. Do these times interfere with production batch cycles, certain online transactions, database backup activity? Is the space management activity being completed on time? Determine the length of time it takes to run space management. Adjust the space management windows to meet all the business needs.

Determine if the space management cycles are spread over multiple hosts or run on the primary host. Be aware of activities being concurrently run at any given time to ensure that they do not cause contention. By determining the location of space management, you can identify any overlap in resources. In most cases, a single primary host is sufficient to execute all space management activities, regardless of how many LPARs are running. There is a limit to the number of tasks that can be executed concurrently.

Be sure to review all of the resources that are being used, including tape drives.

Backup cycles. You can increase efficiency of backup cycles by answering these questions:

- > Does the backup cycle, including the late start time, interfere with any of the items mentioned in the space management discussion? If it does, adjust the times to remove the conflicts.
- > Is the backup cycle run from the primary host or multiple hosts? Your environment determines the most efficient way to run the backup cycle. Look at the number of storage groups, the amount of data that is processed on average and the number of LPARs or HSM tasks that are available.
- > Are the backups being performed on storage groups or pools that do not require HSM backups? Periodically examine which storage groups and pools are being backed up. There's no need to waste precious resources if you don't need to.

SMS constructs. Determine if the data sets are actually going to the proper storage groups to receive the proper HSM management. Review the SMS constructs to determine if the data has the correct management class assigned to the data sets. If the data sets have incorrect management classes, the data could be deleted before its useful life span is used, it could be retained for too long, or it could be in the wrong location for use (such as always on ML2 tape every time the data is needed). All of these situations create unnecessary overhead, headaches for users, and possibly legal issues.

Adding volumes and tape duplexing. To prevent problems when adding volumes, check these items. Does the HSM PARMLIB have the correct exits turned on? Are the proper ADDVOLS coded for non SMS-managed volumes? If incorrect exits are being used, or if the ADDVOLS parameters are coded incorrectly, the volumes will not be added.

For 3590 cartridges that can contain up to 80 GB of data, consider tape duplexing. If an ML2 tape is damaged, trying to recover all of the data could be time consuming. Duplexing is an excellent solution. Be careful when determining the total number of tapes to be used if duplexing because using too many tapes could take all of the available storage capacity for the tape silo. This situation will incur additional storage costs, both robotic and shelf, as well as additional tape media costs. 3590 cartridges can be quite expensive.

Daily activities

As a storage administrator, you are responsible for completing many tasks. Your daily activity may begin with reviewing the error summary reports, activity reports, and thrashing reports.

Check the status of HSM by determining what HSM is enabled to do and what it is actually doing. A good starting point is the volume activity. Understanding how the beginning and ending thresholds look will give you a better idea of what HSM actually processed for each volume. SRM tools enable you to drill down from the volumes to review the actual data migrated.

Check the status of the control data sets by checking HSM utilization and HSM free space.

Check the work done by HSM by looking at what was completed successfully and what was unable to complete. To determine if data should not be migrated, look at the data sets that were recalled. You need to understand who recalled them, when the data was recalled, and how big the data set was. If a large portion of the recalled data sets are small, you should probably leave them on primary disk. In addition, knowing what issued the recall can help you determine if bottlenecks are being created during batch cycles. Taking the recall one step further, the time associated with each recall will determine problems in tape libraries. If the tape comes from outside a library, or if the tape was in use for some other HSM process, the recall could be delayed.

It is important to understand what HSM is processing on a daily basis. This information can help determine where potential problems may lie. You can check the following key items through a combination of queries to HSM and by using an SRM tool:

- > Space migrated
 - Volume report
 - Data set report
- > Error summary report
- > Thrashing reports
- > Recycle by type
 - Zero percent
 - 25 or 30 percent

Divide the recycle of HSM tapes into two groups. Issue the recycle for zero percent. 30 minutes later, issue the recycle for 25 or 30 percent. This process enables you to reduce a larger percentage of tapes in a given period of time. The zero percent tapes do not require a tape mount.

Recycle the tapes based only upon the type of HSM tape. For example, on Monday, issue a recycle only for the Monday Bbackup cycle tapes. Do not include ML2 or any other backup cycle. This process helps reduce the number of tape drives that are used, and it keeps the recycle process on a schedule that is easily tracked.

HSM audit

Performing an HSM audit is slow and painful in most circumstances. Auditing enables you to find, and often correct, discrepancies between control data sets, catalogs, and HSM-owned volumes. For example, a CDS could have hundreds of corrupted records. HSM could continue processing the corrupted CDS for days, weeks, or even months. You may not even realize that a problem exists until a recycle fails or until you try to access the data set with corrupted records. By this time, the damage done could be costly and difficult to repair.

The audit features that are supplied with HSM are difficult to use and can take days to complete. After the audit is started, you should not stop it for any reason because the updates are not processed until after the audit has completed. All updates are nullified if the audit is stopped before completion. A general rule is to run an audit on a monthly basis when other processing is slow. If you are unable to run the audit each month, run it quarterly.

CDS cleanup. HSM uses the backup control data set (BCDS), the migration control data set (MCDS), and the offline (tape) control data set (OCDS). These CDSs have pointers to each other.

The BCDS contains the backup version record (MCC), the data set record (MCB), and the eligible volume record (MCP). The MCDS contains the data set record (MCD), the alias entry record (MCA), and the volume record (MCV).

Good SRM tools can help you determine if there is obsolete data in the CDS. For example, you can see if HSM is backing up ISPF data sets, which is probably not a good idea. These types of files are wasted resources for HSM to manage.

SRM tools can show you other items, such as the oldest data from a migration standpoint in the MCDS, as well as which tape volumes have problems and which ones should be recycled manually to clean the system up.

MAINVIEW SRM REPORTING

Managing HSM is not inherently easy, but SRM tools such as MAINVIEW SRM Reporting can help.

MAINVIEW SRM Reporting makes it easier to work with HSM by providing the following features:

- > real-time HSM messages
- > improved HSM collection database
- > views of the active recall queue within HSM
- > events, TSO messages, and alerts on threshold conditions including queue length, batch recalls that have waited too long, and number of recalls from the same user

- > ability to prioritize HSM recall processing across the Sysplex and HSM's Common Recall Queue (CRQ)

MAINVIEW SRM Reporting provides comprehensive solutions that monitor, analyze, and automate tasks to safeguard the health of storage subsystems and ensure that critical applications complete successfully.

MAINVIEW SRM Reporting provides comprehensive storage information based on real-time and interval-driven monitoring, along with commands and automated tasks. This combination of features enables storage managers to

- > report on storage activities as they occur
- > manage storage subsystem performance
- > observe trend data
- > obtain true minimum and maximum occupancy information
- > establish storage budgets for logical groups and warn administrators or deny allocations when budgets are exceeded
- > obtain application and functional views of storage use
- > quickly obtain the information most commonly used to manage disk resources
- > define thresholds that trigger an unlimited series of self-monitoring, automatic tasks
- > develop customized monitoring and automation procedures for volumes, pools, and logical groups

MAINVIEW SRM Reporting gathers historical information on pools, volumes, data sets, and VTOCs and displays the results. A powerful, user-defined search engine rapidly locates information that is required for daily space management.

Using MAINVIEW SRM Reporting, you can view reports and graphs from an ISPF interface or MAINVIEW Explorer. Daily or monthly trend reports display information about

- > space allocated, used, and idle
- > number of data sets
- > fragmentation index

You can customize batch reports to select fields and the order in which they are displayed. You can generate multiple reports in a single job. Virtually all information can be produced in batch.

Historical performance and capacity information

MAINVIEW SRM Reporting collects detailed historical information about storage performance and capacity. Performance statistics for data sets include response time, I/O rates, and cache activity. Performance and storage-use statistics are collected and summarized at several levels. For example, storage managers can use monthly summaries to forecast storage usage and use daily and interval information to pinpoint a problem. Performance reports are updated automatically at user-defined intervals (snapshots). Real-time activity information is provided for devices, I/O queuing, channels, and contention by enqueue and reserves.

Vendor-specific RAID reports provide data critical to device optimization.

Performance and capacity exception management is a threshold-based facility. For example, storage managers may want to generate an alarm for net capacity load in IBM RVA disk subsystems or a high-cache-read-miss percentage in EMC Symmetrix devices. An alarm can be sent to MAINVIEW SRM Automation, to initiate corrective automated actions.

Application definitions

To meet business needs, storage must be deployed and managed according to its use. Storage occupancy should be viewed from the perspective of logical groupings, such as departments or applications. MAINVIEW SRM Reporting uses *application definition* for these user-defined groupings. DADSM exits (pre- and post-processing exits as well as user-defined exits) enable MAINVIEW SRM Reporting to monitor space at allocation and deallocation to provide an accurate evaluation of an application's storage use versus its quota.

Space availability

Controlling DASD consumption does not usually give the storage administrator authority over users. However, it may be necessary to deny an allocation for a non-critical process, so that business-critical applications have the space they need. The data sets themselves usually are not exclusive to a single job. Therefore, it is important to be able to categorize a data set in various ways and to exercise various kinds of control.

With MAINVIEW SRM Reporting, a data set can be a member of up to four application definitions that can be tiered or hierarchical. Each application definition can have different quota controls that *monitor*, *warn*, or *reject* the allocation if the quota is exceeded. Important applications are assured of space, while those that are optional are restricted. You can remove restrictions when conditions change, such as when the request is made outside peak shift.

Output management

You can automate the response to HSM and DFSMSdss (DSS) messages by automatically generating control statements in response to key HSM errors. Cryptic error log messages are filtered, reworded, and responded to, based on your criteria. This feature saves time and makes HSM easier to manage.

Backup control

MAINVIEW SRM Reporting makes it easy to include or exclude data sets for backup processing. Filtering minimizes coding and ensures that automatic backups do not occur for data sets that do not need them. ALTERD commands are not required.

Recall allocation

MAINVIEW SRM Reporting extends the benefits of HSM to non-SMS-managed data by pooling non-DFSMS data sets that HSM recalls. Within eligible pools, volumes are selected based on the "best fit" of the data set to available extents to control fragmentation.

SUMMARY

Periodically ask yourself the following questions:

- > Is HSM doing what I want?
- > Am I being proactive enough with HSM?
- > Does HSM have internal problems?
- > Is HSM causing outside issues?

MAINVIEW SRM Reporting can help you find the answers. When you can answer these questions, you have a handle on your HSM environment and are able to correct and manage the data center data efficiently and effectively.

Helping you maintain advantage

BMC Software Education Services offers a strategic investment for your business, maximizing the value for your employees and Business Service Management initiatives. Education ensures successful product implementation, promoting mastery of all product capabilities and highest productivity with your BMC Software solutions. To explore our education offerings, visit our web page at <http://www.bmc.com/bmceducation>, or contact BMC Software Education Services by telephone or e-mail:

> **North America**

Telephone: 800 574 4262

E-mail: education@bmc.com

> **Asia Pacific**

Telephone: +61 3 9657 4404

E-mail: ISD_AP@bmc.com

> **Europe, Middle East, and Africa (EMEA)**

Telephone: 00800 26233822

E-mail: emea_education@bmc.com



About BMC Software

BMC Software, Inc. [NYSE:BMC], is a leading provider of enterprise management solutions that empower companies to manage their IT infrastructure from a business perspective. Delivering Business Service Management, BMC Software solutions span enterprise systems, applications, databases, and service management. Founded in 1980, BMC Software has offices worldwide and fiscal 2004 revenues of more than \$1.4 billion. For more information about BMC Software, visit www.bmc.com.



53783