EMC White Paper

# Centera Mainframe HSM Migrator: A DFSMShsm Centera Adapter

***Abstract***

This document describes the Centera Mainframe HSM Migrator for IBM DFSMShsm.

July, 2005

# Table of Contents

# Introduction to DFSMShsm

IBM DFSMShsm® is the leading hierarchical storage manager for MVS storage environments. It provides users and applications with integrated backup, recovery, migration, and space management functionality. The processing by DFSMShsm is controlled by the DFSMS configuration constructs. DFSMShsm provides the following major functions:

Space Management

- Migration – the ability to move unused data sets to lower cost storage such as tape
- Scratch – the ability to remove data sets after expiration or cleanup of temporary datasets
- Space release – the ability to free unused space in a data set

Incremental Backup

- The ability to make logical backups of data sets as they change

Full Volume Dump

- The ability to take full volume physical dumps; these dumps are generally used for disaster backup and recovery

ABARS

- This facility takes logical backups of data sets for disaster backup and recovery

## *EMC Centera Mainframe HSM Migrator*

EMC has integrated the EMC® Centera™ Mainframe HSM Migrator (CMHM) directly into the migration function of DFSMShsm Space Management. This allows administrators to directly utilize Centera as an output device replacing both ML1 DASD and ML2 tape.

The Centera Mainframe HSM Migrator operates within the existing DFSMShsm address space. This results in very few changes to the present operating environment in order to migrate data to Centera. Additionally, migration to Centera is done on a data set by data set basis. However, instead of a data set being migrated to L1 or L2, it is migrated directly to the Centera. Recall for a Centera migrated data set works the same as it does for a data set migrated to L1 or L2. The user or job requesting the data set is unaffected by the fact the data set is being recalled from a Centera. In fact, recall is typically faster because there are no physical tape mounts.

# Centera Mainframe HSM Migrator

## *Version and Platform Requirements*

Operating System Requirements

- IBM z/OS 1.4, with 1.5 or higher preferred
- IBM DFSMShsm License
- One or more Open Systems Adapters (OSAs); GbE highly preferred
- 100 MB of DASD for installation
- DASD for the Centera Mainframe Migrator Control Store

Centera Requirements

- EMC Centera
- CentraStar® 2.4 or 3.0

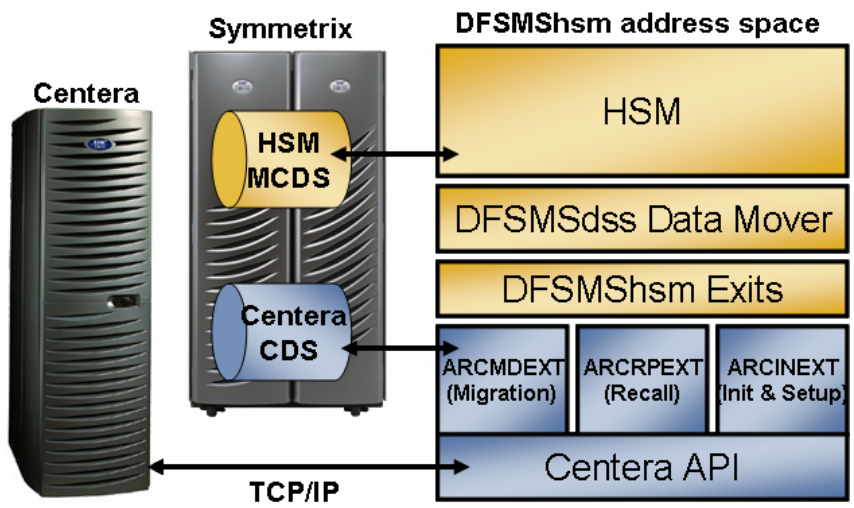
- Centera Mainframe HSM Migrator license

## Installation

Installation of the Centera Mainframe HSM Migrator is simple and straightforward. It assumes that a Centera is available. Installation consists of the following steps:

1. Unload and RECEIVE the XMIT File
2. Extract the Program, Load, and JCL libs (Note: Load libraries must be APF Authorized)
3. Create the Control Store (BLDCS job)
4. Update the CMHM Parameter File
5. Make minor changes to HSM startup procedure to point to:
    - Control Store Data Set (add ECACS DD name)
    - Parameter File (add ECAGCNTL DD name)
    - STEPLIB to the CMHM LINKLIB and IBM C/C++ runtime libraries
6. Add statement to the HSM ARCCMD00 parmlib member to enable exits
    - SETSYS EXITON(IN)
    - SETSYS EXITON(MD)
    - SETSYS EXITON(RP)
    - SETSYS EXITON(SD)
7. Restart HSM

### Architecture Diagram




## Data Movement

The Centera Mainframe HSM Migrator uses the same facility as DFSMShsm to actually move the data, namely DFSMSdss. DFSMSdss has a fully documented and supported interface which allows its use as a generic data mover. During migration, CMHM uses DFSMSdss as the interface to read and format the data set and writes the data sets to the Centera. During recall, CMHM reads data from Centera and uses DFSMSdss to write the data set.

### Supported Data Sets

The Centera Mainframe HSM Migrator supports all NON-VSAM SMS managed data sets except:

- Data sets protected by a RACF Discrete Profile id where the data set RACF indicator is on and in the format 1 DSCB
- Rolled off but not deleted Generation Data Sets (GDGs)
- Extent Reduction Migrations
- Data sets w/explicit expiration dates that have not yet expired

**DFSMShsm User Exits**

The Centera Mainframe HSM Migrator invokes the following user exits:

- ARCINEXT – Initialization and Setup Exit
- ARCMDEXT – Migration Exit
- ARCRPEXT – Recall (and Delete) Exit
- ARCSDEXT – Shutdown Exit (z/OS 1.5 and above)

### Migration

During DFSMShsm Space Management, the ARCMDEXT is called for every data set that DFSMShsm/SMS determines is eligible to be migrated. Then ARCMDEXT decides if the data set should be processed by CMHM. This decision is based on the selection criteria defined in the startup parameters and leverages inclusion and exclusion filters. If the data set is to be processed by CMHM, DFSMSdss is invoked to process the data set.

### Recall

During DFSMShsm Recall, the ARCRPEXT is called for every data set needing to be recalled. Then ARCRPEXT determines if the data set is currently migrated by DFSMShsm or CMHM. If the data set is migrated by Centera, then DFSMSdss is invoked to process the data set.

## *Inclusion and Exclusion Filters*

| Parameter | Description |
|---|---|
| DATASETNAME | Inclusion filter to target data sets to be handled by CMHM |
| MGMTCLASS | Inclusion filter that targets by SMS Management Class name |
| MINIMUMSIZE | Exclusion filter based on data sets' minimum size |
| MAXIMUMSIZE | Exclusion filter based on data sets' maximum size |
| BACKUPPRIOR | Inclusion filter that targets only data sets that have been backed up since last update |

## *Disaster Backup and Recovery*

Disaster backup and recovery is one of the major challenges when working with DFSMShsm.

There are three pieces of data that must reside at an offsite facility:

1. The DFSMShsm Control Data Sets
2. The catalog entries for all migrated data sets
3. The actual migrated data

The actual protection of migrated data is very challenging. Replicating the ever-changing list of migration tapes is a full-time task. The Centera Mainframe HSM Migrator makes this simpler by replicating data sets with Centera Replication. Additionally, the Centera Mainframe HSM Migrator can store its own catalog in the Centera which can also be replicated to a remote location.

### Sysplex Support

The Centera Mainframe HSM Migrator operates in any DFSMShsm environment. It has the same restrictions as DFSMShsm. This means the catalogs, SMS environment, DFSMShsm environment, and GRS environment need to be shared.

### Control Store

The Centera Mainframe HSM Migrator maintains a Control Store which is the rough equivalent of the DFSMShsm MCDS. However, the records within it are far simpler than the DFSMShsm MCDS. In DFSMShsm a NON-VSAM data set migrated to L2 tape has three records in the MCDS and one in the OCDS that must be maintained. Most of the equivalent information for a Centera migrated data set is actually contained in the metadata written to Centera. Only the original data set name, some basic data set information, and the Centera Content Address are kept in a single VSAM record in the Control Store.

### Transparent Operation to Users

The Centera Mainframe HSM Migrator has been designed to be totally transparent to the end user. No external involvement is needed for either Migration or Recall. Migration is fully driven by DFSMShsm and the Centera Mainframe HSM Migrator startup parameters. Recall is fully driven by the normal DFSMShsm recall process.

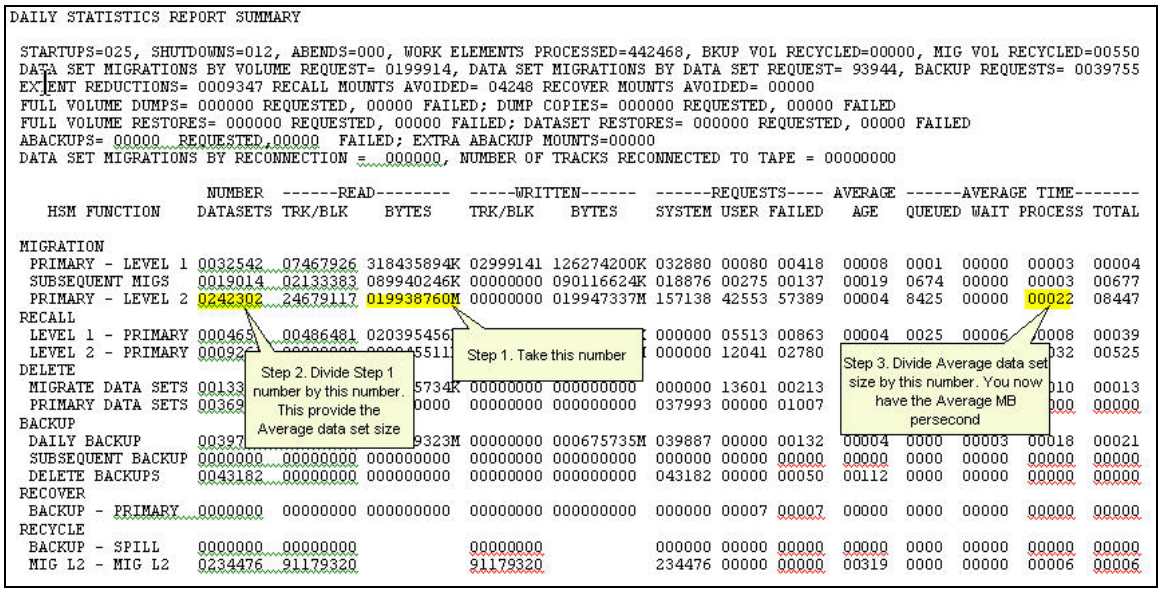
## *Performance*

Sizing for performance is critical in any DFSMShsm environment. The published data transfer rates for a 3590E tape drive is approximately 14MB/s. However, the actual performance obtained by DFSMShsm is considerably less than the published data rates. DFSMShsm has considerable data set level overhead and is typically not able to keep the tape drive continuously busy. Additionally, DFSMShsm does not optimize its write I/O to maximize the data transfer speeds. The current DFSMShsm I/O routines date from the original 3480 and have not been optimized to take advantage of modern tape. Using DFSMShsm's own reports makes it is easy to calculate the actual tape throughput in a system.

### Determining Current Tape Performance

The current tape performance is easy to determine using a DFSMShsm report and doing some simple math. The following procedure can be used to get these numbers:

1. Issue the DFSMShsm command to build the Daily Statistics Report.
2. HSEND REPORT DAILY FROMDATE (yyyy/mm.dd) ODS (data_set_name); see the example report that follows. There is a report for each day since the FROMDATE and a summary for the entire period at the end of the data set.
3. Using the summary, follow the directions in the example. This will provide you with an approximate MB/sec of your current tape system.



```
DAILY STATISTICS REPORT SUMMARY

 STARTUPS=025, SHUTDOWNS=012, ABENDS=000, WORK ELEMENTS PROCESSED=442468, BKUP VOL RECYCLED=00000, MIG VOL RECYCLED=00550
 DATA SET MIGRATIONS BY VOLUME REQUEST= 0199914, DATA SET MIGRATIONS BY DATA SET REQUEST= 93944, BACKUP REQUESTS= 0039755
 EXTENT REDUCTIONS= 0009347 RECALL MOUNTS AVOIDED= 04248 RECOVER MOUNTS AVOIDED= 00000
 FULL VOLUME DUMPS= 000000 REQUESTED, 00000 FAILED; DUMP COPIES= 000000 REQUESTED, 00000 FAILED
 FULL VOLUME RESTORES= 000000 REQUESTED, 00000 FAILED; DATASET RESTORES= 000000 REQUESTED, 00000 FAILED
 ABACKUPS= 00000  REQUESTED,00000  FAILED; EXTRA ABACKUP MOUNTS=00000
 DATA SET MIGRATIONS BY RECONNECTION =  000000, NUMBER OF TRACKS RECONNECTED TO TAPE = 00000000

                     NUMBER  ------READ--------  -----WRITTEN------  ------REQUESTS---- AVERAGE ------AVERAGE TIME-------
    HSM FUNCTION    DATASETS TRK/BLK    BYTES    TRK/BLK    BYTES    SYSTEM USER FAILED   AGE   QUEUED WAIT PROCESS TOTAL

 MIGRATION
  PRIMARY - LEVEL 1 0032542  07467926 318435894K 02999141 126274200K 032880 00080 00418  00008  0001  00000  00003  00004
  SUBSEQUENT MIGS   0019014  02133383 089940246K 00000000 090116624K 018876 00275 00137  00019  0674  00000  00003  00677
  PRIMARY - LEVEL 2 0242302  24679117 019938760M 00000000 019947337M 157138 42553 57389  00004  8425  00000  00022  08447
 RECALL
  LEVEL 1 - PRIMARY 000465   00486481 020395456                      000000 05513 00863  00004  0025  00006  00008  00039
  LEVEL 2 - PRIMARY 000092                                          000000 12041 02780                      00032  00525
 DELETE
  MIGRATE DATA SETS 00133                5734K 00000000 000000000  000000 13601 00213                      00010  00013
  PRIMARY DATA SETS 00369                0000  00000000 000000000  037993 00000 01007                      00000  00000
 BACKUP
  DAILY BACKUP      00397                9323M 00000000 000675735M 039887 00000 00132  00004  0000  00003  00018  00021
  SUBSEQUENT BACKUP 0000000  00000000 000000000 00000000 000000000 000000 00000 00000  00000  0000  00000  00000  00000
  DELETE BACKUPS    0043182  00000000 000000000 00000000 000000000 043182 00000 00050  00112  0000  00000  00000  00000
 RECOVER
  BACKUP - PRIMARY  0000000  00000000 000000000 00000000 000000000 000000 00007 00007  00000  0000  00000  00000  00000
 RECYCLE
  BACKUP - SPILL    0000000  00000000           00000000           000000 00000 00000  00000  0000  00000  00000  00000
  MIG L2 - MIG L2   0234476  91179320           91179320           234476 00000 00000  00319  0000  00000  00006  00006
```



# *Comparing Centera Mainframe HSM Migrator with Traditional Mainframe Storage Methods*

## Comparing Centera to Conventional Tape

On the surface, many people believe that conventional tape is the least expensive option for DFSMShsm. However, the acquisition cost of tape is only a small piece of the overall expense associated with the solution. In today's environment, administrators must consider the cost of automated tape libraries, tape media and maintenance, tape duplexing, offsite vaulting, and migration costs when new tape technology is released.

Tape has a number of significant problems that are unique for DFSMShsm migrated data.

DFSMShsm RECYCLE

- RECYCLE represents one of the biggest management problems for DFSMShsm managed data sets. The RECYCLE process moves valid data sets from existing tapes and places them on new tapes. When a job needs to change a data set that had previously been migrated, the data set is recalled and updated. Depending on the policy, this data set may be re-migrated at a later point in time. Since tape is serial media, not random access, this new copy will go to a different tape. The old copy of the data set is now considered obsolete but still remains in the environment.
- As time passes, tapes contain a significant number of obsolete data sets. The RECYCLE process will move the valid copies of the data to new tapes while the old tapes are returned to the scratch pool. Many sites are forced to run RECYCLE twenty-four hours per day in order to make room for new data sets.
- Beyond the physical impact of RECYCLE, this process is often the chief consumer of MIPS on the mainframe. Using Centera can eliminate the RECYCLE process for obsolete data sets.

When capacity needs to be reclaimed inside Centera, a supplied utility is run as part of a batch job. This utility will "synch" deletes between z/OS and Centera allowing capacity to be reclaimed through a Garbage Collection process in Centera.

### Performance

- Performance in a DFSMShsm environment is critical to servicing requests from users and batch processes. Streaming from tape is often fast once a tape is located, mounted, and positioned. However, getting to the point where a tape is mounted can add significant time to servicing a request. In today's large capacity tape environments, positioning to a specific spot on the tape for recall can take several minutes. Since the Centera acts like DASD tape, mounting and positioning time is not required. This traditionally makes up the majority of elapsed time for a DFSMShsm recall process.

### Reliability

- It is generally accepted that tape media is not as reliable as disk media. In fact, data loss in the modern enterprise due to disk hardware failures is uncommon. Most outages associated with DASD are actually "loss of access" rather than "loss of actual data." While tape reliability has improved over the last few years, tape does not have RAID or other realtime protection mechanisms.
- Virtually all data loss by DFSMShsm involves tape. These failures occur because:
    1) The tape has been overwritten by another MVS program.
    2) The tape media fails.

### Disaster Backup and Recovery

- There are a number of different approaches to handling tape data from a disaster recovery perspective. However, there continues to be substantial challenges to achieving a sound DR strategy in a DFSMShsm environment.
- In many environments, tape data is not considered to have the same service level requirements as "tier 1" data. However, DFSMShsm migrated data sets are considered "online" and a copy is often required at a disaster recovery site. RECYCLE also has considerable impact on disaster recovery. Tape copies must be made when the data is first written, but data is also constantly being moved from tape to tape due to RECYCLE. These new tapes must also be copied to the disaster recovery site.

With Centera Replication, DFSMShsm data sets can be replicated to a disaster recovery site in near real-time over any distance. Centera unidirectional, bi-directional, star, and chain replication topologies are supported.

### DFSMShsm Space Management Multi-tasking

- DFSMShsm multi-tasking is limited to the number of available tape drives. Many sites have difficulty accomplishing their desired Space Management levels because DFSMShsm is unable to accomplish each of its tasks within the allotted window of time. The number one factor limiting DFSMShsm multi-tasking is the availability of tape drives. Because Centera is acting as L1 DASD, this restriction can be lifted.

### Comparing Centera to Virtual Tape

After conventional physical tape, some customers have implemented Virtual Tape. Virtual Tape is the next step-up in the cost matrix and has the following advantages over conventional tape:

- Virtual Tape can present a significantly higher number of tape drives.
- Significantly shorter access time to position the data on a Virtual Tape. This is possible if the Virtual Tape is in the DISK cache. Virtual Tape positioning is simply a matter of calculating the actual position within the file where the Virtual Tape is residing and performing a direct read of that data.
- Actual access to data could be much longer if the virtual volume is not in the cache. This is because the real volume has to be loaded on a tape drive,; the tape must be positioned and the entire Virtual Tape (800 MB+) must be read from tape and written to a disk before DFSMShsm can access the Virtual Tape.

The Virtual Tape advantages over conventional tape come at a higher cost in terms of hardware and management. Since Centera acts like ML1 DASD, it has none of the problems that tape creates for DFSMShsm. In addition to having the same disadvantages as conventional tape, Virtual Tape has a number of its own unique problems:

- DFSMShsm RECYCLE – Since DFSMShsm RECYCLE reads a significant number of tapes during its processing it tends to overrun the cache of the Virtual Tape subsystem. Most sites run RECYCLE very judiciously because of this situation. This results in a significant number of volumes being used by DFSMShsm in the Virtual Tape subsystem as well as the physical tapes that back the Virtual Tapes.
- Disaster Backup and Recovery – Virtual Tapes add another level of complexity to the disaster recovery environment. Virtual Tapes are eventually written to proprietary tapes. These generally cannot be copied or sent to a recovery site. The Virtual Tapes must either be copied to a real tape or the Virtual Tape subsystem must somehow replicate the Virtual Tape to a remote Virtual Tape subsystem located in a facility off site.

### *Comparing Centera to ML1 DASD*

Finally, ML1 DASD is an option for DFSMShsm migration. Its advantages over tape are:

- No RECYCLE
- No limits placed on DFSMShsm multi-tasking
- DASD is the most reliable media for data storage

However, there are several disadvantages:

- The cost of ESCON- or FICON-attached disk is traditionally expensive
- The maximum size of a migrated data set is limited to a single volume

Utilizing channel attached DASD for DFSMShsm migration is extremely costly when compared to using Centera. Additionally, there is no limitation in data set size when migrating data to Centera.

## Conclusions

Centera Mainframe HSM Migrator has many advantages over current DFSMShsm migration target devices:

- Roughly the equivalent of DFSMShsm Level 1 DASD except no size limit on data set
- Much lower hardware acquisition cost than traditional z/OS DASD
- RECYCLE is no longer necessary
- No physical tape mounts
- Simple disaster recovery
- DFSMShsm multi-tasking restrictions are eliminated
- Data sets move directly to Centera; no movement from Level 1 to Level 2