Systems

Introduction to the IBM 3850 Mass Storage System (MSS)

IBM

Systems

Introduction to the IBM 3850 Mass Storage System (MSS)

IBM

Preface

This publication is designed to convey information about the IBM 3850 Mass Storage System. It is intended for data processing executives, operations managers, systems programmers and other professional data processing personnel. It assumes familiarity with data processing concepts and terminology.

Contents

Figures

IBM System/370 Model 158 and the IBM 3850 Mass Storage System (MSS)

Introduction

Direct Access Storage Devices (DASD) have steadily grown in capacity and speed, but only data with a relatively high value and use can be maintained on DASD permanently, for two reasons. First, when we calculate the cost per megabyte of DASD storage (including the cost of drives, control units and packs), it is substantially higher than the corresponding cost for tape. Second, for many users, the total volume of data collected and maintained cannot fit on even the maximum configuration of DASD. For these reasons, the data processing community has come to recognize the need for an economical mass storage device.

Tape systems also have grown in terms of capacity (recording densities) and in speeds. However, three characteristics of tape devices limit their usefulness. First, data stored on tape is inherently sequential, so that random or direct accessing is impractical. Second, tape volumes are typically not mounted until they are called for, whereas most DASD packs tend to remain more or less permanently mounted. This causes human intervention, which implies both a time delay in mounting the tape and a larger possibility for human error than is typically encountered with DASD. Third, usually only one data set is stored on each reel of tape. An entire, unsharable device is required for each mounted reel of tape, regardless of the activity rate. This can result in low utilization of tape drives and it adds to scheduling complexity.

It would be beneficial, in most installations, to have another alternative for data storage—a mass storage system. Such a system should have:

- Capacity equivalent to a tape library.

- Availability and mounting of volumes under control of the operating system.

- Data organizable in the variety of methods available on DASD.

- Data transfer rates approximately equal to DASD.

- Cost per megabyte of data storage closer to tape costs than to disk costs.

The 3850 Mass Storage System (MSS) is IBM's response to these requirements.

Consider another trend in data processing: with the advent of larger memories, the user can sustain a higher level of multiprogramming than before. That is, more tasks can be handled concurrently.

Each task—whether a batch job step or initiated by a terminal user—requires access to *data sets*. These data sets reside on *volumes*, which must be mounted on I/O *units*. Increasing the number of concurrent tasks increases the requirements for available units, volumes, and data sets. Thus, as the number of concurrent tasks supported by a system increases, the data *requirement* of the system increases.

In the past, as a system's requirement for data has increased, the only solution has been to attach more tape drives and disk drives to the system. However, the additional drives—along with the additional tape reels and disk packs required by these drives—generate operational problems, such as:

- The library storage and retrieval procedures for tape reels and disk packs grow larger and more cumbersome. The time for a volume to be located in the library increases and the time for a volume to cycle from demount to the

library and become available for re-use grows significantly longer. (Already, in some installations, this time is as long as 24 hours.)

- Security of volumes in the computer room usually is sacrificed for speed of recycling through the library.

- Operators must mount/demount more volumes on more devices, increasing the time between mount request and mount complete, and increasing the probability of mount errors and subsequent rerun/recovery costs.

- The costs of the I/O devices increase and the requirement for floor space in the library and the computer room grows larger.

Once the additional I/O units are installed in a complex environment to support more concurrent tasks, it is difficult to maintain balanced systems utilization. There will be occasions when even more devices are needed, while at other times not all available units will be required. The MSS addresses these problems—satisfying user requirements for an economical, large capacity storage device with a sophisticated technology that extends the concepts of virtual storage to the I/O components of a computer system.

The Mass Storage System—An Overview

The IBM 3850 Mass Storage System consists of IBM 3330 Disk Storage and one or two IBM 3851 Mass Storage Facilities.

The MSS provides very large data capacity, all under system control. The capacity range of MSS is from 35 billion (35 x 10^9) bytes to 472 billion (472 x 10^9) bytes (4,720 3336-1 volumes).

Data under control of the MSS is stored in DASD format images on low cost media (magnetic tape in data cartridges) until it is needed. The MSS transfers data from the data cartridges to DASD (3330 drives), when it has been requested, in a process called *staging*. Once data has been staged, it behaves precisely like any other data resident on a 3330 in terms of organization and accessibility. When the data is no longer needed, and the space it occupies on DASD is required for other data, any "cylinder" containing new or updated data is *destaged* back onto the data cartridge (Figure 1).
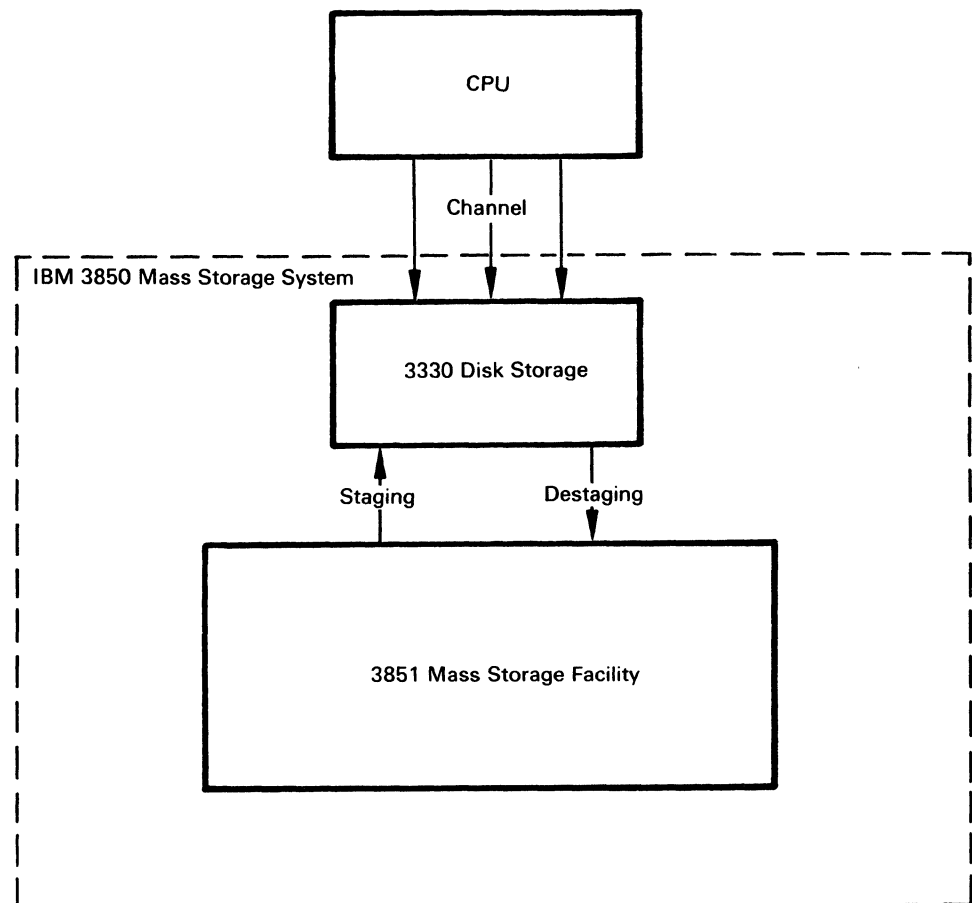
Figure 1. Mass Storage System Data Flow

All cartridges under control of the system are physically within the 3851 Mass Storage Facility (MSF) of the MSS. Floor space for storage media is greatly reduced. Many tape data sets may now be stored as DASD data sets, reducing or even eliminating tape drive and tape reel requirements.

Each data cartridge can hold up to 50 megabytes of data, so two cartridges are required to store the data capacity of an entire 3336 Model 1 disk pack (Figure 2). Indeed, two cartridges are always initialized together and the pair is referred to as a *mass storage volume* (MSV) and the MSV is managed by the Mass Storage System. 3336 Model 11 disk packs contain the equivalent of two MSVs.
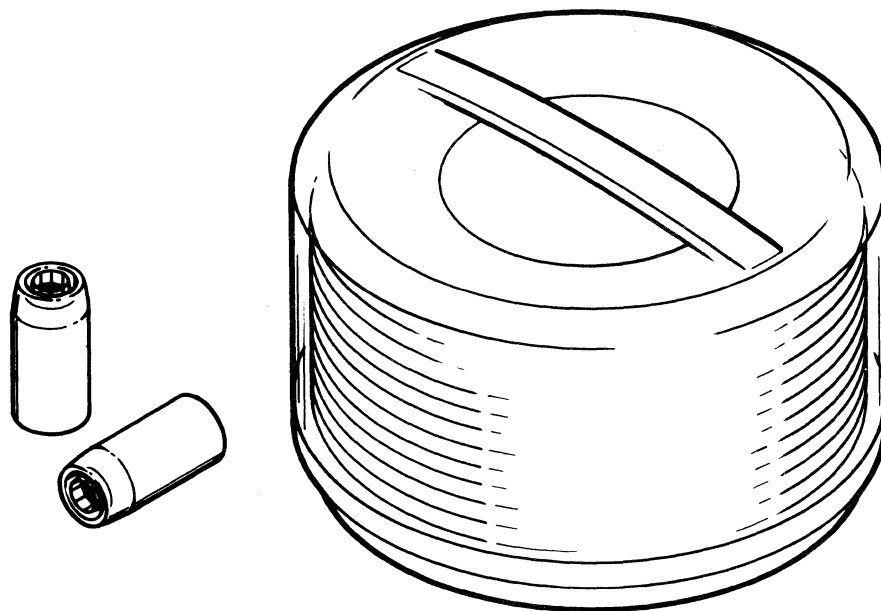
1 Mass Storage Volume =2 Cartridges=1 3336 Model 1 Disk Pack=100 Million Bytes

Figure 2. Mass Storage Volume Capacity

Because MSVs when staged appear as normal DASD volumes to OS/VS, the operating system issues mounts for these volumes onto DASD drives. However, only the portions of a volume that are requested will be staged onto DASD. The MSS hardware has been designed to present to OS/VS the image of having more DASD devices available than are physically present (the *virtual device* concept). OS/VS and user programs request data as before, and the hardware translates these requests to obtain data from its actual location on physical DASD or in the Mass Storage Facility.

This eliminates the time to manually pick tape reels from a library and greatly reduces mount times. The cycle time of a volume through a library system is dramatically shortened, because **all** volumes are stored and retrieved under control of the system.

To assist in managing up to 4,720 volumes, IBM provides new operating system functions and utilities. These routines help convert data sets to mass storage volumes, keep track of space availability, group mass storage volumes for simplicity of management, and create backup.

The MSS is a hierarchical storage mechanism that combines many of the advantages of tape and disk systems with potential for operational benefits in a complex environment (Figure 3).
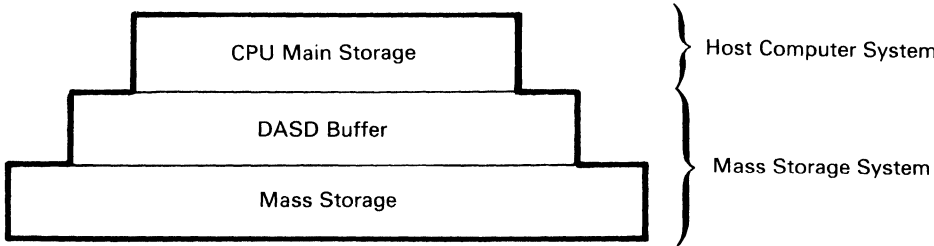
Figure 3. Mass Storage System Hierarchical Store

Data is contained in the Mass Storage Facility, the DASD buffer, or main storage as it is required, and migrates up and down in the hierarchy as its use dictates.

The Mass Storage System—Hardware Functional Description

The major components of the Mass Storage System are (Figure 4):

- The 3851 Mass Storage Facility (MSF)
- The 3830 Model 3 Storage Control
- The 3333 Disk Storage and Control (Models 1 or 11)
- 3330 Disk Storage Drives (Models 1, 2, or 11).
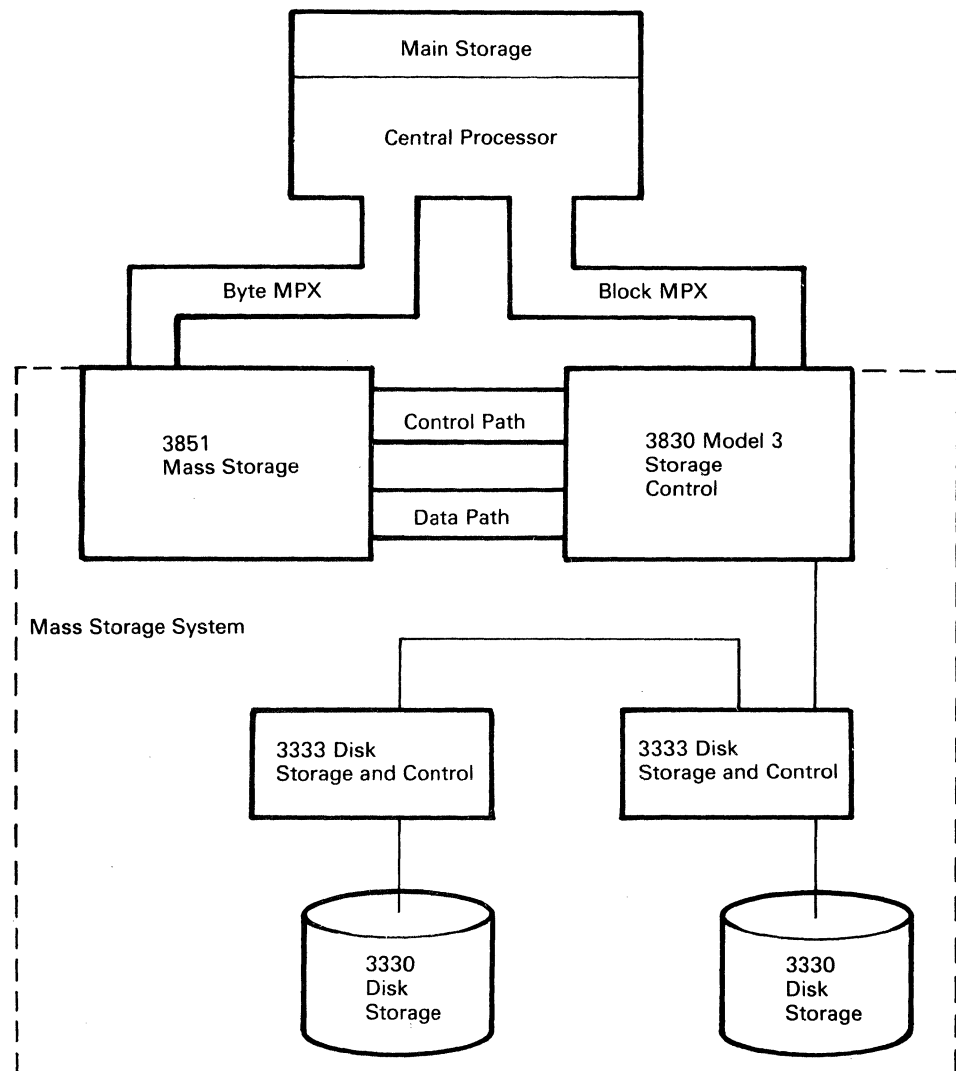
Figure 4. Mass Storage System Configuration

Optionally, the System/370 Integrated Storage Control feature on the Models 158 and 168 can be equipped with a Staging Adapter. The Staging Adapter provides identical function to two 3830 Model 3 Storage Controls except for the number of channels that can transmit data to host processor(s). When a Staging Adapter is installed on an Integrated Storage Control, the entire ISC (both sides)

is equipped to operate in the MSS. One half of the ISC cannot be so equipped without the other. In this publication references to the 3830 Model 3 Storage Control also apply to an Integrated Storage Control with the Staging Adapter.

3851 Mass Storage Facility (MSF)

Each 3851 Mass Storage Facility contains up to 236 billion (236 x 10^9) bytes of data and moves data to and from DASD as needed. Two 3851 MSFs can be included within the same 3850 Mass Storage System.

The 3851 Mass Storage Facility includes these major components:

A. Data cartridges to store the data in the storage cells

B. Data Recording Controls (DRCs) and Data Recording Devices (DRDs)—to transfer data between the MSF and the 3830 Storage Control

C. Cartridge Access Station—to accept data cartridges into or remove them from the MSF

D. Accessors to retrieve and deliver data cartridges and accessor controls to control the accessor movements

E. Mass Storage Control—to provide overall control of MSS functions

Most of the 3851 Mass Storage Facility components are shown in Figure 5. Data is stored on data cartridges that reside in the cartridge storage cells. When data is requested, the accessor moves to the appropriate storage cell, selects the cartridge from the cell, and transports it to a data recording device. The cartridge is then loaded into the DRD, which reads data from the tape and transmits it to a 3830 Storage Control through the data recording control.
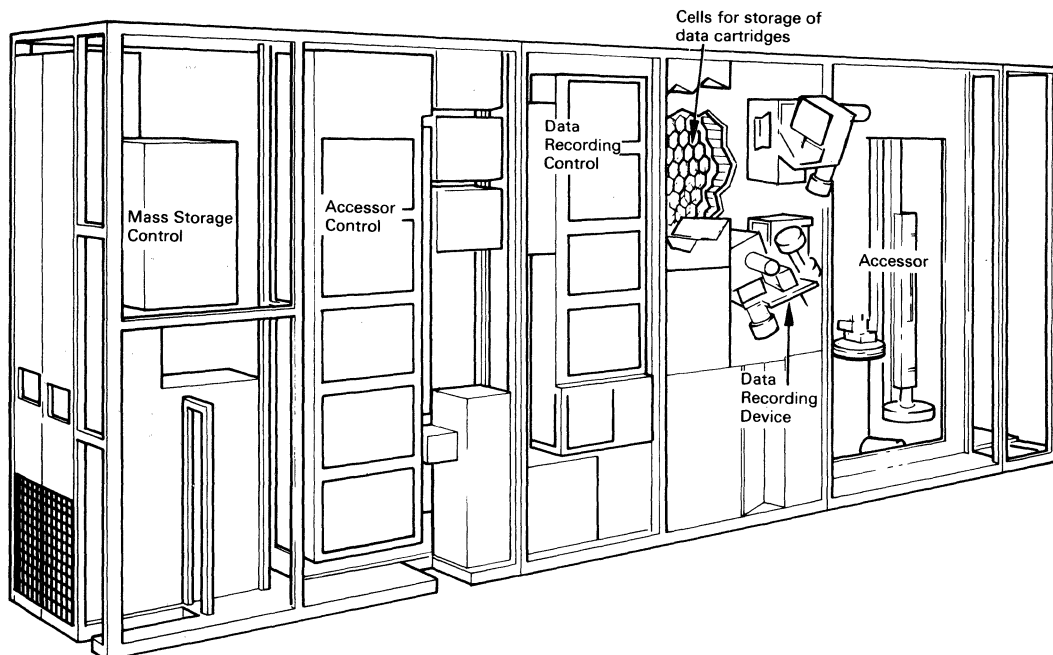
Figure 5. 3851 Mass Storage Facility Components

When the required data has been completely staged, the data cartridge is returned to its home storage cell. The staging/destaging operations are supervised and directed by the Mass Storage Control.

A. Data Cartridge

The 3850 Mass Storage Facility uses data cartridges to store data. Each cartridge can contain up to 50.4 million bytes of data; the cartridge is approximately 2 inches (5.08 cm) in diameter and 4 inches (10.16 cm) long. The cartridge holds a spool of tape 770 inches (17.5m) long (Figure 6).
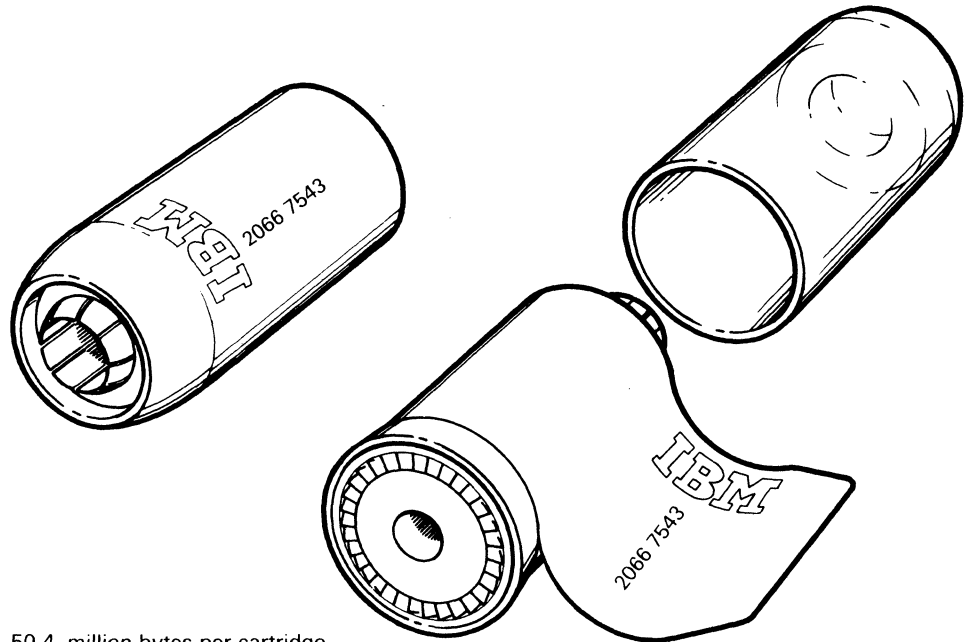

```
50.4  million bytes per cartridge
 770  total inches of tape
 677  inches of usable tape
   2  inches diameter
   4  inches length
```

Figure 6. Data Cartridge

Data is recorded on the magnetic tape in the image of 3330 Disk Storage cylinders. Each cylinder is recorded on a fixed location on the tape, and specific cylinders can be located by unique identifiers along the edge of the tape. Each cartridge is capable of storing 202 cylinders in the 3330 format. Therefore, two cartridges are the equivalent of one 3336 Model 1 Disk Pack.

The *mass storage volume* is the storage reference for data sets within the Mass Storage System, just as tape volume and disk pack volume are the storage references for tape and disk data sets.

Mass storage volumes within the 3851 Mass Storage Facility have the following characteristics:

404 cylinders

19 tracks per cylinder

13,030 bytes per track maximum

100 megabytes

B. Data Recording Device

The 3851 Mass Storage Facility includes one to four pairs of data recording devices (DRDs) and one data recording control (DRC) for each DRD pair for reading and recording data from and onto cartridges. The DRD also has a high speed search capability that makes storing multiple data sets on one cartridge practical.

C. Cartridge Access Station

The 3851 Mass Storage Facility has a cartridge access station which allows manual entry and removal of cartridges. There are separate ports for entry and exit of cartridges.

D. Accessors

Two accessors and associated accessor controls and power supplies are included in every model of the 3851 Mass Storage Facility. The accessors provide for the movement of data cartridges in the Mass Storage Facility; that is, from a storage cell to a data recording device and back.

E. The Mass Storage Control (MSC)

The Mass Storage Control (MSC) in the 3851 Mass Storage Facility controls all staging and destaging operations. The MSC's functions include:

1. Accepting requests for data from up to four System/370 CPUs, Models 145, 155II, 158, 165II, 168, 158 multiprocessor and 168 multiprocessor.

2. Determining the location of data by referring to an inventory list of cartridges and mass storage volumes.

3. Allocating space in eight-cylinder increments on staging DASD for data to be staged.

4. Instructing the accessor controls to move a cartridge containing the requested data from its storage cell to a data recording device.

5. Initiating the staging of the data from the cartridge in the DRD to the allocated space on disk storage.

6. Performing error recovery procedures, alternate path retry, and device reallocation as needed during the staging/destaging operation.

7. Monitoring the amount of allocable disk storage space. When the amount of allocable space becomes less than a specified threshold (or when the host system so instructs), initiating deallocation and destaging of changed cylinders to create additional allocable space available for other data set requests. The criteria for selection of cylinders to be destaged is based on a least-recently-used (LRU) algorithm.

8. Maintaining usage and error statistics by component and cartridge.

9. Recording the physical configuration of the MSS:

   • Data and control paths

   • Status of components

   • Automatic switching of components

10. Maintaining a record of the location, attributes, and status of all mass storage volumes.

As OS/VS controls the movement of data between 3330 Disk Storage and the CPU, the Mass Storage Control controls the movement of data between Disk Storage and the Mass Storage Facility.

F. Model and Feature Options of the 3851 Mass Storage Facility

There are eight models of the 3851 Mass Storage Facility; they differ in (1) size and (2) the number of MSCs they have.

The following table shows the differences among models:

	Models A1,B1*	Models A2,B2*	Models A3,B3*	Models A4,B4*
Billion bytes (10^9) of storage capacity	35	102	169	236
Maximum Number of Data Cartridges	706	2,044	3,382	4,720
Mass Storage Volumes	353	1,022	1,691	2,360
Data Recording Controls	1	2	3	4
Data Recording Devices	2	4	6	8
System 370 Channel Interfaces with Two Channel Switch, Additional	4	4	4	4
MSC Ports with MSC Twin Port Feature**	2	2	2	2

* "A" Models have one MSC each; "B" Models have two (one is for backup).

** One MSC Port allows the MSC to attach to eight Mass Storage System elements.

An *element* is a 3851 Mass Storage Facility, a 3830 Model 3 Storage Control or an Integrated Storage Control with the Staging Adapter (each ISC counts as two elements). The maximum configuration with one MSF consists of the MSF and seven 3830 Model 3 Storage Controls (eight if the Twin Port feature is installed). The maximum configuration with two MSFs consists of the two MSFs and fourteen 3830 Model 3 Storage Controls. When the MSS configuration includes two MSFs, each staging drive must have a path to each MSF. Paths to an MSF can be provided by attaching the 3830 to each MSF or by attaching two 3830s (one to each MSF) addressing a common set of 3333s via the String Switch. This effectively limits the number of staging drives in a configuration with two MSFs to 128. A minimum MSS configuration consists of an MSF Model A-1, one 3830 Model 3 Storage Control (or an Integrated Storage Control with Staging Adapter), and two 3333 Disk Storage and Controls (must have four disk drives).

Each model can be equipped with two special features:

1. **Two Channel Switch, Additional**, which expands the number of System/370 channel interfaces on the MSC from the standard two to four. With this feature, the MSC can attach to a maximum of four central processors or a maximum of two multiprocessing systems.

2. **MSC Twin Port**, required to increase the allowable number of Storage Controls to 8 in a single MSF configuration and to 14 in a configuration that has two MSFs (an ISC counts as 2 Storage Controls).

G. Attachment

One or two 3851 Mass Storage Facilities can be attached to IBM System 370 Models 145, 155-II, 158, 165-II and 168, 158 multiprocessor or 168 multiprocessor. If two MSFs are included in an MSS configuration, both must be "A" Models. Both Mass Storage Facilities operate under the control of one Mass Storage Control. The other MSC provides the backup function identical to that in the "B" series. A maximum of two multiprocessor systems may be attached to an MSS.

3830 Model 3 Storage Control

The 3830 Model 3 Storage Control or the Staging Adapter on an Integrated Storage Control Feature performs the following functions:

• Transfers data between disk drives and the CPU, via the 3333.

• Maps virtual device address and logical (cylinder, head and record) address to physical device address and actual (cylinder, head and record) address. For more information, see "Theory of Operations."

• Detects a request for data that is not present on DASD (cylinder fault).

• Requests the MSC to resolve a cylinder fault.

• Transfers data between data recording drives via a data recording control and staging disk drives.

A buffer is provided within the 3830 Model 3 Storage Control to overlap data transfers. Data staged from and destaged to the Mass Storage Facility first goes to the buffer in the 3830 Model 3. While this data is being written from or into the buffer, the 3830 Model 3 can simultaneously transfer data to or from a System/370 CPU.

The 3830 Model 3 has two channel interfaces. One channel interface is used by the Mass Storage Control. The other is used for data transfer to the CPU. Two additional CPU channel interfaces are possible with the addition of the special feature (#8171) Two Channel Switch, Additional. The Integrated Storage Control on the System/370 Models 158 and 168 with the Staging Adapter has two interfaces per path. One interface is used for communication with the Mass Storage Control and the other is for data transfer to the CPU. One additional channel interface is possible per path with the addition of special feature (#7905) Two Channel Switch for ISC.

Up to thirty-two 3330-type drives can be attached to one 3830 Model 3; however, only sixteen 3330 Model 1 or eight 3330 Model 11 drives can be used as staging drives. (The Mass Storage Control considers a 3330 Model 11 Drive as two 3330 Model 1 drives when it is used as a staging drive.)

The 3830 Model 3 Storage Control can attach to a maximum of four data recording controls in a Mass Storage Facility. Each path of an ISC with the Staging Adapter can be similarly attached.

The 3830 Model 3 or ISC with the Staging Adapter has an expanded addressing capability: up to 64 unique addresses per channel interface (maximum of 192 per 3830-3 or 128 per Integrated Storage Control path) are possible. This expanded addressing capability is necessary so that more virtual volumes than physical drives can be simultaneously mounted. (See "Theory of Operations.")

The 3333/3330 units function exactly as before.

The Mass Storage System—Theory of Operations

In order to better understand the operation of the Mass Storage System, keep in mind that the space on all staging drives under control of the MSS is used as buffer storage between the Mass Storage Facility and the CPU. This space is managed by the Mass Storage Control.

Staging Units

Some of the disk drives connected to a 3830-3 can be designated for use as buffers (in which case they are called by any of the names: *staging units, staging devices, staging DASD* or *staging drives*), while others may be designated as standard disk drives (thus, *non-staging units, devices, DASD*, or *drives*).

Although 3336-1s or 3336-11s can be used on a staging drive (thus becoming *staging packs*), virtual volumes (data staged from mass storage volumes) are always in the format of 3336-1 (100 megabyte) packs. Staging packs must be formatted in a special manner:

* The volume serial numbers of all staging packs have the same first four characters (user chosen). The last two characters of the volume serial number must be the MSS identification number.

* The VTOC must be on cylinder zero, track two.

* The VTOC indicates that the entire pack is filled by a single data set (hence, OS/VS cannot allocate space on this pack, which has the effect of reserving the pack for space allocation by the Mass Storage Control).

The format of a staging pack is such that

* A non-staging pack **cannot** operate on a staging drive without being reinitialized.

* A staging pack **cannot** operate on a non-staging drive without being reinitialized.

Virtual Drive Concept

The implementation of virtual drives is one of the most significant concepts in the MSS. With the MSS creating the image of many 3330 drives, the system is able to satisfy more concurrent requests for devices and volumes and should enable more jobs to run at the same time than with conventional tape/DASD.

Virtual drives are implemented at the 3830-3 level. The standard unit addressing consists of three hexidecimal digits (12 bits). The first four bits designate the channel number, which remains unchanged. However, the *physical* characteristics of the system from the channel level down and assignment of the other eight unit-address bits are:

* A maximum of four 3830-3s can be attached to a single channel, so it takes two bits to select which 3830-3 is being addressed on the channel.

* A maximum of four 3333s can be attached to a particular 3830-3, so it takes two more bits to designate which 3333 is to be selected on the 3830-3.

- A maximum of eight direct access storage drives can be connected to a particular 3333, so it takes 3 bits to specify which drive is being selected.

Thus, to actually designate which drive is being addressed from a channel takes 7 bits. But there are eight bits in the non-channel portion of the device address. Mathematically, this provides a *virtual* addressing capability on each channel interface on a 3830-3 of 64 units—regardless of whether there are only 2 drives or the maximum of 32 attached to the 3830-3 (Figure 7).
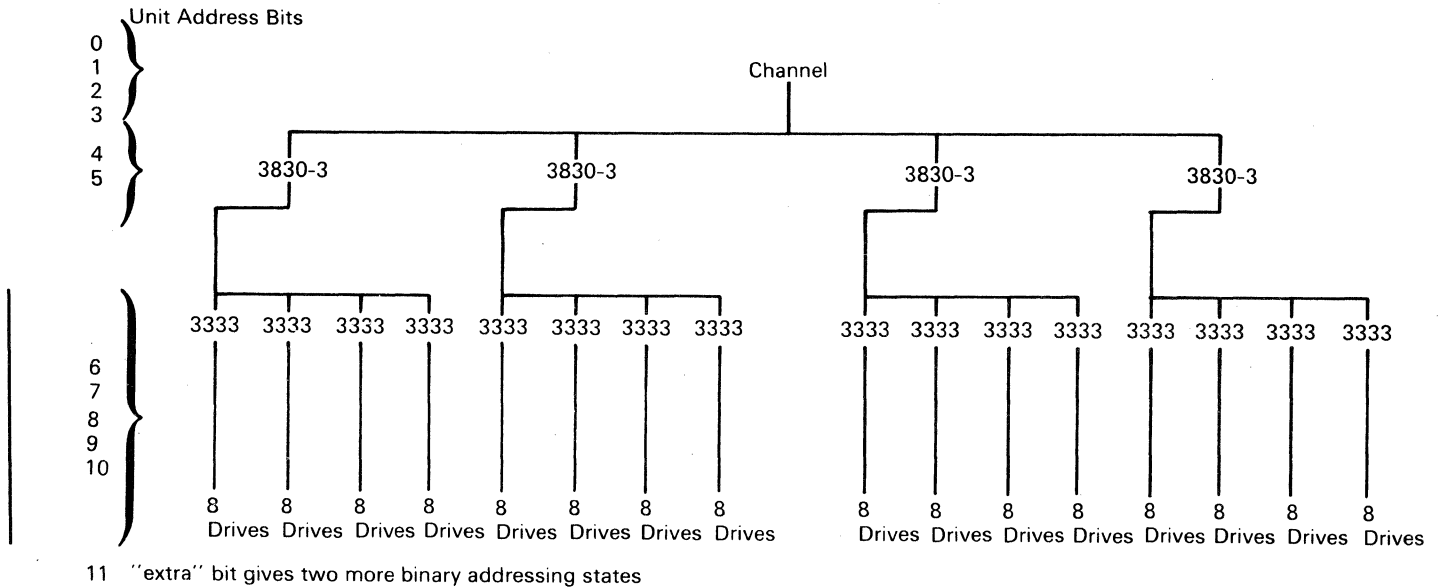
Each 3830-3 can address 4x8x2=64 units. Only 32 can be physical; hence, "virtual units".

Figure 7. Virtual Unit Addressing

Suppose, for example, on block multiplexer channel "2" we have a 3830-3 designated as number zero (B"00"). Then the range of virtual addresses on this 3830-3 for this channel interface is "200" to "23F". We describe this entire range of addresses to OS/VS as 3330 type unit control blocks (UCBs) with a bit turned on to designate the drive as a virtual unit. A request to device "21F", will be sent to channel "2" and the first 3830-3 (number B"00"). The 3830-3, then, will translate the "1F" portion of the address to the appropriate location on a staging drive under its control.

A virtual unit, then, is a concept in the same sense as virtual address space in OS/VS. A virtual unit is conceptually divided into pages of eight contiguous cylinders in the same manner a virtual address space is broken up into 64K segments. DASD space to hold a page of a virtual volume can be allocated from **any** available pages of the staging drives controlled by a particular 3830-3. It is possible that one page of a virtual volume is located on one staging pack, while the adjacent page of the same virtual volume is located on a different staging pack. However, both pages will have the same virtual unit address.

The status and location of pages, cylinders, virtual drives, virtual volumes, and staging drives are maintained in tables in the MSC and the 3830-3. While the MSC allocates space on the staging drives in eight cylinder pages, data is staged and destaged in increments of one cylinder.

All the staging drives under a 3830-3 can logically be divided into two groups of distinct physical units (staging drive groups) for work-balancing purposes. This enables page allocations to be balanced between the staging drive groups under a 3830 Model 3. All pages from a particular mass storage volume, mounted on a given virtual unit, must reside within the same staging drive group. This activity is managed by the MSC.

A Typical Job

With this background, let's follow an example of a job using data sets on mass storage volumes. The allocation routines of OS/VS recognize a request for a virtual volume (and a corresponding virtual unit upon which to mount the volume) through new DD parameters in JCL. The new parameters signal allocation to request assistance from the MSS software support.

When an existing data set is requested (DISP=OLD or DISP=SHR), the MSS locates the data cartridges containing the mass storage volume, allocates a page of staging space, and stages up cylinder 0 of the requested volume. When the volume verification routine of the Mount command reads the volume serial, the request goes through the 3830-3 which translates the virtual device address and virtual volume location from cylinder 0, track 0, record 3 ('00003') to the device and location on that device where that record has been staged. To the host system, everything appears as conventional DASD.

When a request for a new data set is made and a volume serial is specified, the process is the same, except that after volume verification, the DADSM (Direct Access Device Space Management) routines read the VTOC to locate space for the new data set. If the VTOC is not on cylinder 0, the 3830-3 recognizes that the cylinder requested has not been staged up—a *cylinder fault* occurs (refer to "Glossary" for definition of cylinder fault). The MSC and 3830-3 together find another page of staging space, locate the cartridge containing the requested cylinder, and stage the requested cylinder. Then the request is satisfied. The appearance to the system is that the "seek" channel command took longer than normal to complete.

When a request for a new data set is made, but no volume serial is specified, the selection is first directed to the Mass Storage Volume Control (MSVC) functions of the Mass Storage System Communicator (See "Program Support"). After assistance from MSVC, the allocation request is handled as though it is a request for a specific volume. At this point the virtual volume has been mounted on a virtual unit which has been assigned to a staging drive by the MSC. Sometime during the execution of the program, the user will issue OPEN to these data sets. At OPEN time, the MSS is requested to allocate enough pages of staging DASD to hold the entire data set and, if this is an existing data set, to stage the entire data set. (**Exception:** ISAM data sets are not staged at OPEN—see "Program Support" for ISAM). OPEN will not wait for completion of data set staging before returning to the user program. If a **new** data set is OPENed, staging space is allocated, but no staging takes place.

The user program eventually issues an I/O request to an open data set on a virtual volume. The request goes through the access method modules it has always gone through. These modules convert the request into system terms–control blocks and channel commands–as before. A request for a record identifies a unit address and the cylinder, track, and record number of that record. The request goes through IOS exactly as before. When the channel program reaches the 3830-3, this device translates the virtual unit address to the actual address of the staging device that contains the requested page and the

cylinder number of the virtual volume to the cylinder of the staging drive containing that virtual cylinder's data. Because data is staged and destaged in increments of cylinders, the track number and record number do not need to be translated. The I/O operation is completed in the usual manner.

A request for cylinders not yet staged results in a *cylinder fault*. A cylinder fault may occur because the data set was not completely staged before the request for this record was made. The 3830-3 notifies the MSC, and works with the MSC, to locate and stage the desired cylinder. If the data set is still being staged, cylinder faults will be resolved after the staging operation completes. The I/O operation then is completed normally. The appearance to the system is that the "seek" command took longer than normal to complete. Note that ISAM data sets **always** operate in cylinder fault mode (See "Program Support").

Control of Staging Space

Recall that the MSC and the 3830-3 maintain tables that reflect the status of all that is happening in the MSS. One of these tables maps virtual volume pages to staging DASD pages and the entries in this table reflect, for each cylinder in a given page, whether some record on that cylinder has been referred to and whether it has been modified. Whenever a cylinder is referred to, the table entry describing the page containing that cylinder is given a timestamp. As the system goes about its work, a shortage of staging space may be encountered (that is, the amount of allocable space becomes less than a user-chosen lower threshold). When this occurs, the MSC must free some of the occupied space in that staging drive group.

To free space, the MSC examines the timestamp associated with each page. The least-recently-used (LRU) pages are selected, until enough pages have been selected to bring the available staging space above a user-chosen upper threshold or until there are no further pages old enough to be considered for selection. (The user defines how many timestamps ago constitutes "old enough"). In all pages thus selected, cylinders that have been modified are destaged, and the pages are relinquished (made available for subsequent allocations). The MSC keeps track of what data was on each page until the page must be reallocated. Thus, if a request for a destaged data set occurs, the current copy may still exist intact on staging DASD. *Data Reuse* routines in OS/VS attempt to locate data in this manner first, to prevent unnecessary staging. This is analogous to page reclamation in virtual storage management.

When the problem program finishes with a data set, it issues a **CLOSE** macro. **CLOSE** operates as before. Usually, a data set on a virtual volume is not immediately destaged upon **CLOSE**; the data remains on staging DASD until the LRU algorithm needs to free the staging space containing that data. Users of VSAM data sets can request to wait until the data set is destaged before control is returned from **CLOSE**. In this case, the staging space is immediately queued for destaging. This ensures that, if an I/O error occurs during destage, the user is notified while his program is still in execution. (For more detail on the operation of common access methods see "Program Support".)

A virtual volume remains mounted on a virtual device until OS/VS or the console operator demounts the volume. When a demount request is received, all staging pages containing data from the virtual volume currently mounted are scheduled to be destaged (if required) and reallocated. Only those cylinders which have been modified are actually destaged.

Tape Recording Organization

Understanding how the MSS stages and destages data requires understanding how data is organized on data cartridges.

The first step to this is an awareness of how tape is threaded onto the data recording devices (DRDs). The tape is wrapped around a read/write mandrel in a helix-like position (Figure 8). The read and write heads revolve within the mandrel. If one were to take the tape out of its cartridge and lay it flat, he would see that the resulting recording pattern is a diagonal *stripe* (Figure 9).
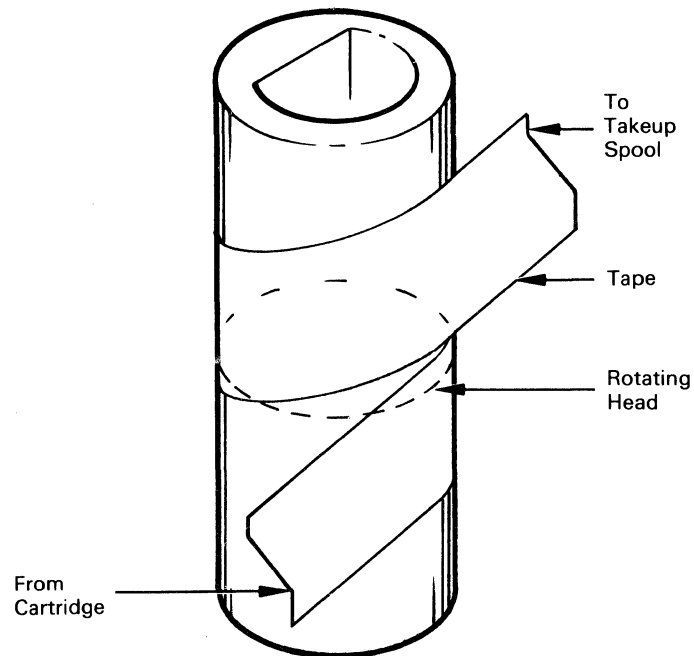
Figure 8. Tape Threaded in Recording Position
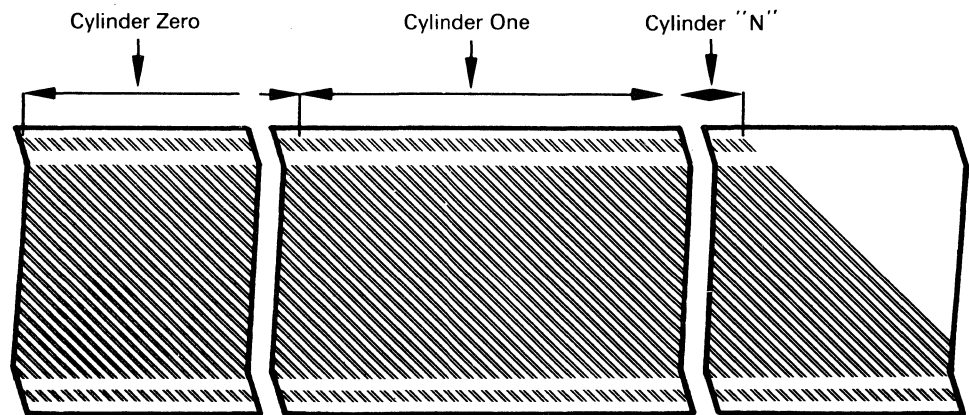
Figure 9. Stripe Format

The stripes on the tape are numbered starting with zero. The first ten stripes contain the cartridge label area. There is room for 4096 data bytes per stripe. There are 67 stripes allocated per 3336-1 cylinder image. Four of these stripes are alternates and one is a separator. Each stripe carries its own identification or stripe number.

Servo information along the outside edges of the tape assists in positioning the tape properly in the DRD. Included in this information is the stripe number, which assists in the high-speed search for a cylinder on a volume.

Each group of 67 stripes corresponds to a particular cylinder address on a virtual volume enabling a search of stripe addresses to locate a selected cylinder. Since data is always staged and destaged in cylinder increments, data is recorded sequentially from the first track in the cylinder to the first stripe of the 67 stripes assigned to that cylinder. An end of track indication is recorded by the hardware as this condition is encountered. Count, key (if any), and data are recorded, but home address, gaps, and DASD ECC information are not.

Mass Storage System Program Support

This chapter describes OS/VS System Control Program (SCP) support and problem program use of the 3850 Mass Storage System.

Overview of MSS Support Components

A primary function of the 3850 Mass Storage System is to store large amounts of data under "system control". This data is staged when needed and destaged to a data cartridge when no longer needed. Once staged, the data appears to the application programmer as it normally appears on an IBM 3330 DASD device.

The primary additions to OS/VS provide a *virtual unit* and *virtual volume* appearance to the problem program and the system operations personnel. The virtual unit concept permits an OS/VS operating system to use more "drives" than actually exist in its DASD hardware configuration. OS/VS maintains a virtual unit control block for each MSS virtual DASD unit. These control blocks appear to represent physical DASD drives, but they do not. There should be more virtual unit control blocks defined at systems generation time for OS/VS than there are physical DASD drives in the host's I/O configuration.

The virtual volume concept permits many partial volumes to reside on a single MSS staging drive. It also permits different portions of a mass storage volume to reside on several staging drives. In the MSS environment, the mapping of staged data is performed entirely outside the host operating system, and is transparent to the host. For example, data from cylinder 8 of a mass storage volume may be staged to cylinder 25 on a staging drive without host operating system translation or knowledge.

The primary support components for the 3850 Mass Storage System are illustrated in Figure 10. This figure defines three major areas where MSS support resides in addition to the problem program area. They are:

1. OS/VS area

2. Mass Storage Control (MSC)

3. Special DASD data sets.

These areas and the problem program area are briefly summarized below.

Problem Program Area

The problem program area in the System/370 CPU is used for execution of user application programs and for the following OS/VS components:

- MSC Table Create Program (MSCTC)

- OS/VS System Generation (SYSGEN)

- OS/VS Utilities

- OS/VS Job Management routines

- OS/VS Data Management routines

A new program called *Mass Storage Control Table Create* must be used to define the MSS processing environment prior to any use of the MSS. This program creates a set of tables that contain MSS configuration information and other
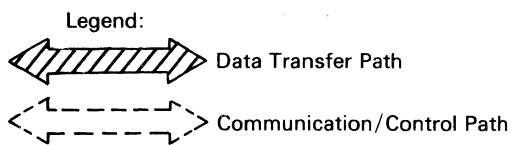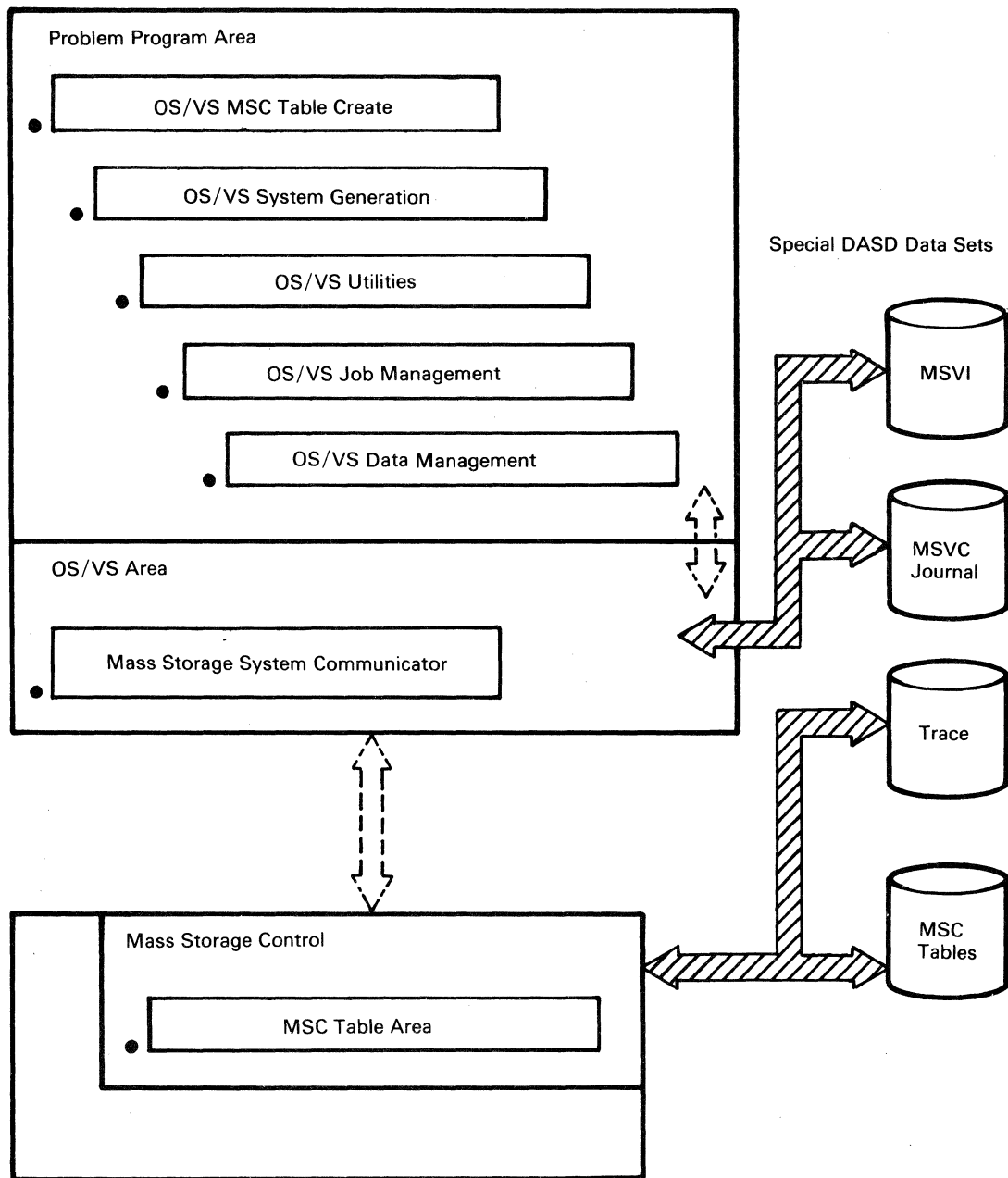
Figure 10. Primary OS/VS Support Components for the IBM 3850 Mass Storage System.

control information. These tables are then written on two disk packs that reside on specific 3333/3330 disk drives. These tables occupy approximately 32 cylinders, the remaining space on each pack being available for staged data. These tables are accessed and updated by the Mass Storage Control. The program prints card images and optionally punches cards (IODEVICE and UNITNAME control statements) to be used in the OS/VS system generation.

The *OS/VS system generation* process required to support the MSS will involve the specification of some additional stage I parameters unique to the MSS. The card images which are produced by the MSCTC program will also become part of the OS/VS system generation input stream.

The *OS/VS utilities* that support the MSS operating environment can be classified as follows:

- Access Method Services Utilities

- DASD Utilities

- Recovery Utilities

- Serviceability Utilities

These utility programs are tools to assist users (for example, system programmers, system operators, and MSS space managers) in the management of an orderly MSS operating environment.

OS/VS Job Management has been enhanced to support mass storage volume selection for jobs that process MSS-resident data sets. These enhancements allow the user to specify new JCL DD statement parameters and issue new operator commands which are unique to the MSS environment. The Allocation sub-component of Job Management requests the mounting and demounting of mass storage volumes on virtual units. These mount/demount requests are directed to the MSC and are not sent to an operator's console. As a result, these requests do not require manual intervention by operations personnel.

OS/VS Data Management provides the access methods which are used by application programs to access data sets. The programs that process MSS-resident data sets must function as though those data sets were resident on IBM 3330 DASD devices. This allows user programs to use any access method appropriate for DASD data sets. Such access methods include VSAM, QSAM, BSAM, BPAM and BDAM (ISAM is described later in this section). Device-independent programs that process tape resident data sets via QSAM may need only minor modifications to process the same data sets in the MSS.

OS/VS Area

The OS/VS area of the System 370 host operating system contains a new component called the *Mass Storage System Communicator* (MSSC) that provides support for MSS functions.

The MSSC is the interface between the System/370 host operating system and the Mass Storage Control in the 3851 Mass Storage Facility. All host programming requests for such functions as the mounting and demounting of virtual volumes, and staging and destaging of required data sets are communicated to the MSC via the MSSC component. Likewise, all MSC-originated communications with the System/370 host are directed to the MSSC component.

A set of functions within the MSSC helps keep track of the status of mass storage volumes and assists in the selection of volumes for new data set allocation. Collectively, these are called Mass Storage Volume Control (MSVC) functions. They are automatically integrated into the OS/VS System Control Program during system generation of MSS support. MSVC maintains a record of the amount of free space available on each Mass Storage Volume, updating its records each time space is allocated. MSVC also records the mount status of MSVs, updating its records at volume MOUNT and DEMOUNT time.

The MSVC functions provide access to a pair of special DASD data sets: the Mass Storage Volume Inventory (MSVI) data set, and the MSVC Journal data set. These data sets contain information regarding mass storage volumes that reside in the 3851 Mass Storage Facility.

Mass Storage Control

The MSC contains the microprogrammed logic and a record of the hardware interconnections required to manage all activity within the 3850 Mass Storage System. To do this, the MSC must have access to tables of information initially created by the MSCTC program. These tables reside on two staging drives and are continually updated by the MSC.

If the alternate MSC (available on "B" models of the 3851 MSF and dual "A" model configurations) is invoked by the host(s), it restores the tables and processing continues.

Special DASD Data Sets

There are four primary DASD data sets that support the Mass Storage System. Two are needed to support the Mass Storage Volume Control functions and two are required to support the Mass Storage Control functions.

The *MSVI* data set contains control records for each mass storage volume in the installation. These mass storage volumes may be resident in the 3851 MSF or currently stored outside the MSF (in archival shelf storage, for example). These control records are updated by the host MSSC as a mass storage volume's characteristics change.

The *MSVC Journal* data set is used to record all modifications to *MSVI*. This data set is updated whenever the MSS Access Methods Services utilities are executed. If access to the *MSVI* data set becomes impossible, the recovery procedure involves updating an earlier copy of the *MSVI* data set, using the current *MSVC Journal* data set as input.

The *Trace* data set is created on user request. It contains time-stamped activity records which reflect data cartridge movement within the MSF, and staging and destaging activity within the MSS. Reports which use the Trace data set as input can be used in MSS performance analysis, system tuning, job scheduling and load balancing.

The *MSC Tables* are required by the MSC to control such MSS activity as staging and destaging data. The required tables are loaded (from DASD) into MSC storage during the Initial Microprogram Load (IML) of the 3851 MSF. The MSC updates its tables internally and on DASD, as required, to reflect the current status of all activity in the MSS.

Mass Storage Control Table Create (MSCTC)

MSC Table Create builds the MSC tables on 3336 Models 1 or 11 disk packs. The content of the tables is determined from user-supplied control statements that define the planned configuration of:

- CPU(s)

- Mass Storage Facility(s)

- 3830 Model 3 Storage Control Unit(s), and/or Integrated Storage Control features

- 3330 DASD.

The MSC tables contain information about the MSS that is needed by the Mass Storage Control to manage the staging and destaging of data within the subsystem. These tables reside on each of two staging drives, which must be on separate 3333 strings.

The MSC tables contain four basic classifications of data:

- Configuration information, which describes the hardware present in the Mass Storage System and the interconnections between the various hardware components.

- Cartridge information, which describes the locations and volume serial numbers of the mass storage volumes in the library.

- Activity information, which describes the status of staging and destaging of data.

- Control information, which contains information required to locate all the tables and also the basic information required to complete the Initial Microprogram Load (IML).

The MSC updates cartridge, control, and activity information as required, to reflect the current status of the MSS.

Input and Outputs

The MSC Table Create program uses the following input:

- A control data set, which contains utility control statements that define the MSS hardware configuration and identify staging drive groups.

- An input MSC Tables data set, which is required for all executions of MSCTC after the first one.

The MSC Table Create program produces the following output:

- A new MSC Tables data set.

- Punched system generation IODEVICE and UNITNAME control statements to be used in Stage I of system generation for the OS/VS host operating system(s) (optional).

- A message data set, which contains all messages produced during the execution of MSC Table Create.

- Two reports that are also contained in the message data set:

  1. Installation configuration map describing all physical connections of the CPU and MSS hardware components.

  2. Diagnostic messages.

The technical details for using MSCTC are described in the publication *OS/VS Mass Storage Control (MSC) Table Create*. Figure 11 shows the MSC Table Create process.

OS/VS SYSGEN

Following the successful execution of MSCTC, the user can proceed to generate the OS/VS Systems Control Program in the normal fashion. This would involve selection of additional Stage I parameters related to MSS, and the inclusion of the UNITNAME and IODEVICE statements punched by MSCTC. VSAM support must be included in the system generation process.
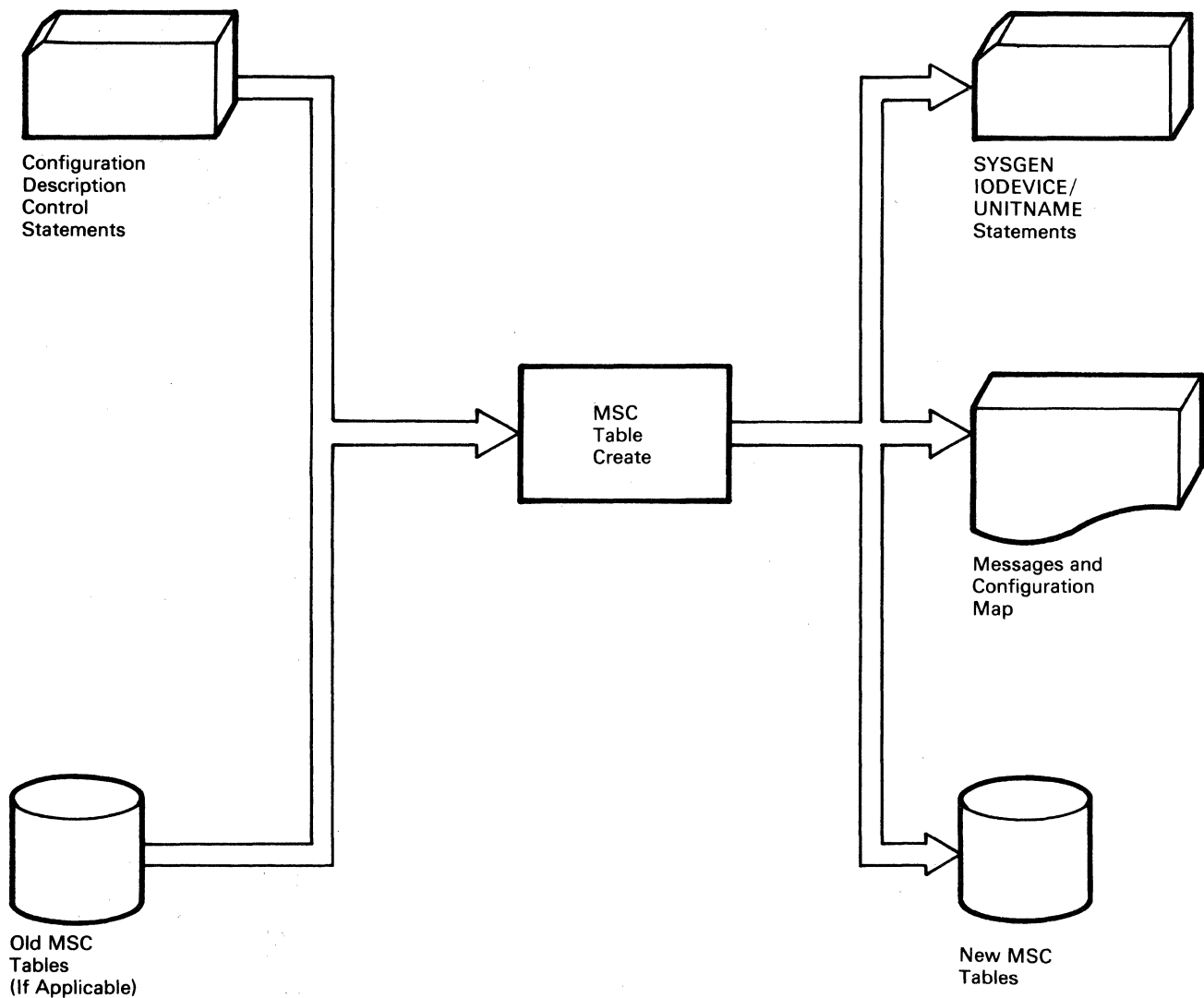
Figure 11. Mass Storage Control Table Create

Mass Storage System Communicator (MSSC)

The MSSC is an OS/VS component which provides the communication between a System/370 host operating system and the MSC portion of the 3851 Mass Storage Facility.

All host programming, such as OS/VS Job Management, OS/VS Utilities, and OS/VS Data Management that require MSS functions will invoke the MSSC component which subsequently issues the appropriate command(s) to the Mass Storage Control. Figure 12 shows this MSSC communication flow. Some commonly requested MSS functions include:

- Mounting or demounting of a mass storage volume
- Staging or destaging of an application data set
- Ejecting a mass storage volume from the 3851 MSF.
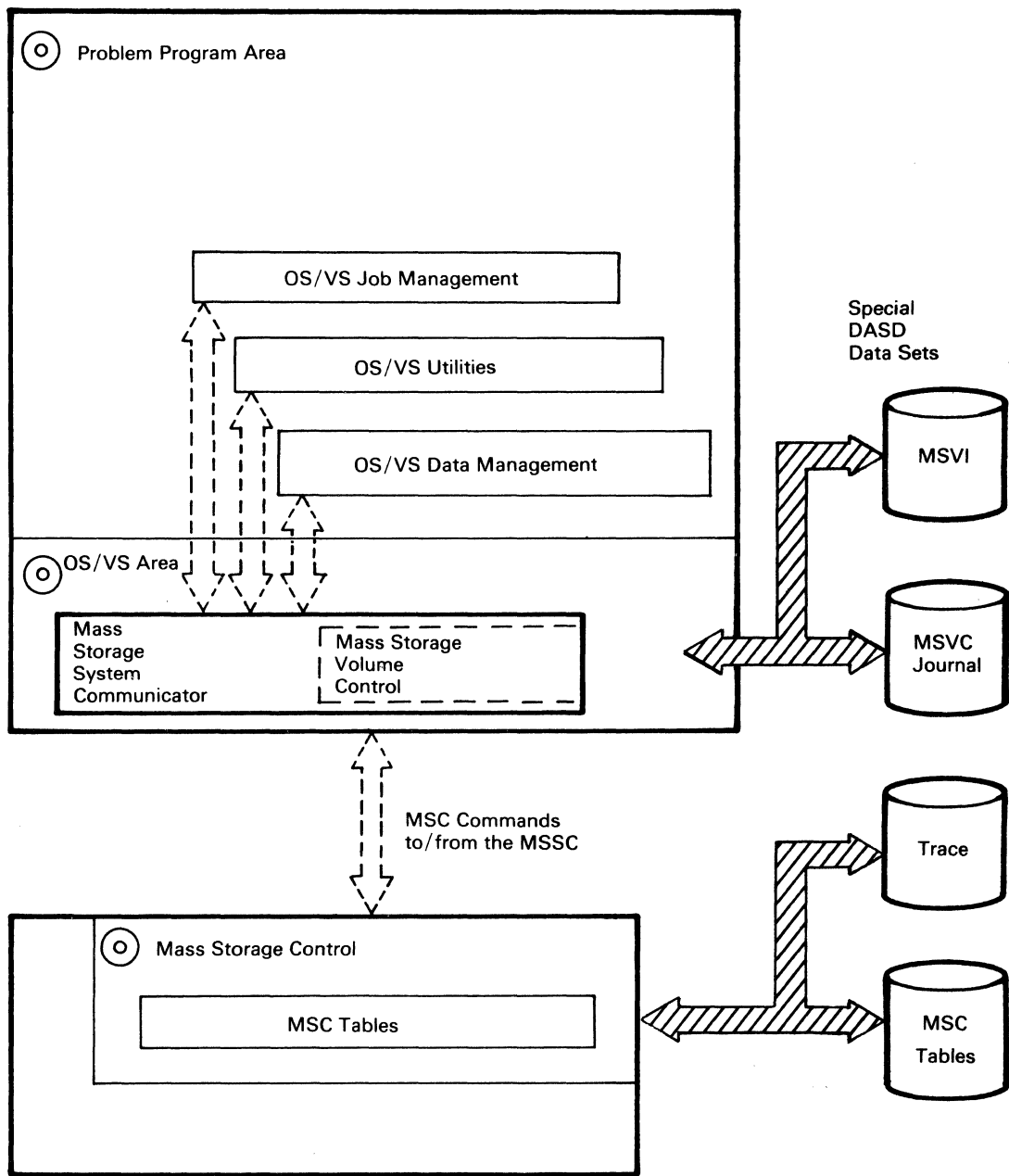
Figure 12. Mass Storage System Communicator Flow

The Mass Storage Volume Control functions of the MSSC provide access to information about mass storage volumes. This information, which is centralized on DASD and shareable between hosts, facilitates the management of the mass storage volumes which can reside in the MSS. The information maintained for a mass storage volume includes:

- Volume serial number

- Volume owner's identification

- Data cartridge serial numbers that contain the volume

- Amount of unallocated space on the volume.

Let's now examine the MSSC and its relationship with the MSC.

MSSC/MSC-Related Functions

Commands that the MSSC issues to the MSC originate from four host programming facilities:

- OS/VS Job Management routines

- OS/VS Utilities

- OS/VS Data Management routines

- OS/VS IPL/NIP (Nucleus Initialization Program)

The following paragraphs describe these facilities in regard to the MSC commands that the MSSC issues on their behalf.

OS/VS Job Management-Related MSC Commands

The Job Management functions that must be translated into MSC commands by the MSSC include:

- *Mount/Demount,* to request that a virtual volume be mounted on a virtual unit (drive) prior to being accessed by a problem program, and to request that a virtual volume be demounted from a virtual unit (drive) when access to a virtual volume is no longer required.

- *Vary On/Vary off,* to request that the MSC start or stop using a specific device or component of the MSS.

- *Purge/Assign Primary Host,* to request that all mounted volumes for a particular host be demounted, or to request (via Assign) that any unsolicited MSC messages be directed to a particular host in a multi-host environment.

- *Suspend,* to indicate that the host wishes to quiesce MSS processing for a power down, or to inform the MSC that the host MSSC will no longer access the MSS.

The mount/demount functions can be invoked from the allocation routines, the termination routines, or operator command processor routines of Job Management. All other functions are invoked via the operator command processor routines. See the section in this chapter entitled "User Programs" for further details about Job Management.

OS/VS Utility-Related MSC Commands

There are new functions available with Access Method Services Utilities that provide for the definition, creation, and management of mass storage volumes in the MSF. These functions require the services of the MSC and are translated into MSC commands by the MSSC. They are:

- *Define Volume* or *Modify Volume (DEFINEV, MODIFYV)*, to request that a new mass storage volume be created using two "scratch" data cartridges, or that the attributes or characteristics of an existing mass storage volume should be altered.

| - *Move Volume* or *Cartridge* or *Convert Volume (EJECTC, EJECTV, CONVERTV)*, to request that a single data cartridge or a mass storage volume be physically ejected from the MSF through the cartridge access station.

| - *Copy Volume/Copy Cartridge (CREATEV, COPYV)*, to request (via Convert Volume) that a "real" 3336 Model I DASD volume be copied to/from a mass storage volume, that a mass storage volume be copied to another mass storage volume, or to request (via REPLACEC) that one cartridge in the MSF be copied to a different cartridge.

There are other new Access Method Services utility functions that provide a way to manipulate certain MSC hardware functions. These functions are usually invoked by systems programming personnel. These functions are translated into MSC commands by the MSSC, and conveyed directly to the MSC. They are:

- *TRACE,* to request that cartridge movement activity and data stage/destage activity be monitored and recorded in the MSC trace tables for future access by the host.

- *TUNE,* to request that the current LRU tuning parameters in the MSC be displayed or altered.

All of the new Access Method Services utility functions are described in more detail under "OS/VS Utilities" later in this chapter.

OS/VS Data Management-Related MSC Commands

The OS/VS Data Management functions that can be translated into MSC commands by the MSSC are:

- *Mount/Demount,* to request that a mass storage volume be mounted on a virtual unit (drive) prior to being accessed by a problem program, and to request that a virtual volume be demounted from a virtual unit (drive) when access to it is no longer required.

- *Acquire,* to request that staging DASD space be allocated prior to the staging of data set extents, and optionally, to request that those data extents should be staged.

- *Relinquish,* to request that the staging space bound to a staged data set be unbound, and optionally, that cylinders containing new or modified data records be destaged to their appropriate data cartridges.

The Acquire function must be accomplished by the MSC before the application program can successfully access data sets that reside on virtual volumes. The OPEN routines for each OS/VS access method (that is, VSAM, QSAM, BDAM, etc.) will issue the appropriate Acquire request to the MSSC at the time a particular data set is OPENed by a problem program. The options that can be specified for the Acquire function depend upon which OS/VS access method is being used to access the data set.

·The Relinquish function is optionally invoked by the close routines of the Virtual Storage Access Method (VSAM) at the time a VSAM data set is closed by a problem program. The other OS/VS access methods do not request the Relinquish function in their respective CLOSE routines.

For more details on OS/VS access method support of the MSS, see "User Program" in this chapter.

All of the MSSC commands to the MSC outlined above cause the MSC to update its tables. Some of the commands require access to the MSVI data set.

OS/VS IPL/NIP-Related MSC Commands

The Purge, Initialize, Assign Primary Host, and Ready functions are completed during IPL/NIP to initialize the Host-MSC interface.

Let's now examine how the MSSC (via MSVC functions) manages the Mass Storage Volume Inventory.

Mass Storage Volume Control Functions

The Mass Storage Volume Control (MSVC) functions of the Mass Storage System Communicator provide a mechanism to aid in managing mass storage volumes in the MSS and to ease the conversion from tape. Because of the Mass Storage System's capacity, centralized control of mass storage volumes is a necessity. The MSVC functions maintain information for each mass storage volume which exists in an installation. This information is kept in *volume records* on a system data set called the Mass Storage Volume Inventory (or Inventory) data set.

Besides maintaining information for individual mass storage volumes, the MSVC functions allow a user to identify and control several mass storage volumes (a logical group) that share common characteristics. This facility, known as mass storage volume *grouping*, can be implemented via Access Method Services utility functions (see "OS/VS Utilities"). After the user defines a mass storage volume group and its characteristics, he assigns individual mass storage volumes to the group.

Information about mass storage volume groups is maintained in a *group record* by MSVC on the Inventory data set.

The user defines a mass storage volume group to facilitate management of data sets and free space that exists on volumes assigned to that group. For each group, the user can specify the following characteristics:

- *Primary* and *Secondary Space Allocation Defaults* to be used during the allocation of new data sets to a volume within a group. This can eliminate the need for space parameters during the creation of a new data set.

- *Free Space Threshold* that is used to monitor the amount of free space for volumes assigned to the group. When the amount of free space for the group drops below the threshold, the user is notified.

- *Release Option,* also an allocation default, which causes the release of allocated but unused space during allocation of new data sets to the group. This can eliminate the use of the release parameter during the creation of data sets.

- *Retention Period* is the user-specified period of time to retain a volume. It is used to generate an expiration date for all volumes in the group based on the date each volume is first used by MSVC.

Other group parameters allow you to specify description, owner, cartridge location, bind/no bind, exclusive/share, DASDERASE/no DASDERASE, and read-only/read and write access.

Data sets are allocated to volumes that have later expiration dates than the expiration date on the data set.

The facility to group volumes is provided to assist the application programmer and the MSS space manager.

Assume that the following group has been created.

Symbolic Groupname	MSV Serial Numbers Assigned to the Group	Purpose of This Group
PROLLGRP	111111 222222 333333 444444	To contain data sets that are related to one application—PAYROLL.

When the application programmer needs to create a new data set containing payroll information, he simply replaces the volume and space parameters on the DD card with a new parameter that specifies the name of the group—PROLLGRP.

During data set allocation for this new data set, OS/VS Job Management and the MSVC functions cooperate in finding a volume on which the data set will reside, relieving the user of the requirement to specify space parameters. The presence of the new parameter causes the allocation routines to invoke the MSSC. The MSVC functions of the MSSC then find a volume within PROLLGRP that contains enough space for the data set, for example volume 111111. The data set is allocated to volume 111111. After volume 111111 is demounted, MSVC will update the *volume record* and the *group record* to reflect the amount of remaining free space on the volume 111111, and the group PROLLGRP. If the free space threshold of the group has been reached, the MSS space manager is notified.

Several grouping techniques can be employed to control the allocation of data sets. Groups can be established to contain data sets that:

- are related to a specific application

- are used by a specific user department

- are approximately the same size.

The Mass Storage Volume Inventory is a VSAM, key-sequenced data set shareable between hosts in a multi-host environment. The MSVC Journal data set is a sequential data set that is also shareable. These two data sets do not have to reside on a mass storage volume, as they can be on DASD external to the MSS. For Inventory data set updates, the MSVC will write a control record to the Inventory Journal data set. If the Inventory data set is destroyed, the Journal records can be used to restore it by applying them against a backup copy. The Inventory Journal data set is also a repository for messages which MSVC wishes to bring to the attention of the MSS space manager. These messages must be extracted and reviewed periodically.

The MSVI is updated when the status of a mass storage volume changes. Examples of such status changes include:

- *Creating* a mass storage volume, which involves initializing two new, or "scratch", data cartridges.

- *Copying* a mass storage volume (for backup purposes) to another mass storage volume.

- *Storing* a mass storage volume, which designates a volume as non-mountable to a host CPU, even though it may remain inside the MSF.

- *Ejecting* a mass storage volume, which involves the physical ejection of a non-mountable mass storage volume from the MSF.

These mass storage volume status changes are invoked by the MSS space manager via Access Method Services Utilities and will result in the MSC tables and the Inventory data set being updated to reflect the changes.

There are other mass storage volume status changes that are **not** initiated via OS/VS Utilities, but must be reflected in the MSC tables and Inventory data set. Examples of these changes include:

- *Mounting* a mass storage volume, as a result of a host Job Management MOUNT request, before staging of any data from the volume.

- *Demounting* a mass storage volume, resulting from a host Job Management DEMOUNT request, which occurs when the volume is no longer required by user programs.

- *Inserting* a mass storage volume through the cartridge access station.

The host MSSC, via its MSVC functions, has ultimate responsibility for maintaining Mass Storage Volume Inventory information regarding individual mass storage volumes and groups of mass storage volumes.

Having examined the System Generation requirements for MSS, and the communication between the System/370 Host operating system and the MSC, let us now focus on the OS/VS Utilities that provide extensive support of the MSS.

OS/VS Utilities

This section describes the major utility functions that support MSS. They can be classified in four categories: Access Method Services Utilities, DASD Utilities, Recovery Utilities, and Service Utilities.

These OS/VS Utility functions assist user personnel, such as system operators, MSS space managers, and system programmers, and IBM support representatives to:

- Initialize

- Maintain

- Tune

- Measure

- Service

the 3850 MSS.

Access Method Services Utilities

The use of the Access Method Services utility functions is mandatory to support the MSS. The new functions are implemented via new and modified utility commands specified in utility control statements. These commands are unique to the MSS environment and can be classified as follows:

- Cartridge management commands

- Volume management commands

- Group management commands

- Report generation commands

- MSC control commands

- VSAM data set definition commands

The first four classifications are provided to assist those persons designated as "space managers" to effectively use the data storage capacity of the Mass Storage System.

Before examining the cartridge management commands, let's examine the process of inserting cartridges into the MSF.

Manual Cartridge Entry

Entry of data cartridges into the MSF is accomplished by simply placing a cartridge into the cartridge access station. When this is done, the MSF automatically selects the cartridge from the cartridge access station loads the cartridge into a DRD, and reads the cartridge label. If this cartridge is not part of a mass storage volume, it is considered to be a "scratch" cartridge and available for subsequent use as part of a mass storage volume. All **new** data cartridges are considered scratch cartridges. If the cartridge is a scratch, it is then delivered to an empty MSF cell, and an MSC table (Scratch Cartridge List) is updated to reflect the availability and location of the scratch cartridge.

If the cartridge is part of a mass storage volume (not a scratch cartridge), then the cartridge is delivered to an empty MSF cell and an MSC table (Transient Volume List) is updated to reflect the location and identity (mass storage volume serial number and cartridge serial number) of the data cartridge. The MSC will also send an unsolicited message to the MSSC in the primary host CPU. This message contains information that identifies the new cartridge as part of a mass storage volume. The information is then used to update the MSVI data set by the MSSC, so that the Inventory data set will reflect the physical presence of the new cartridge now located in the MSF.

Cartridge Management Commands

The cartridge management commands are:

- EJECTC

- REPLACEC

The EJECTC command physically ejects scratch cartridges from a particular MSF. As each cartridge is ejected, the EJECTC command will issue a message to the operator identifying the cartridge. At completion of the utility operation, EJECTC prints the cartridge serial numbers of all ejected cartridges.

The REPLACEC command replaces one cartridge of an active mass storage volume with another cartridge. The cartridge to be replaced may be a defective cartridge or one becoming defective. The SYS1.LOGREC data set identifies

defective cartridges (cartridges to which one or more cylinders could not be destaged).

Mass Storage Volume Management Commands

Mass storage volume management commands are:

- ADDV
- CONVERTV
- COPYV
- CREATEV
- EJECTV
- MODIFYV
- RECOVERV
- REMOVEVR
- SCRATCHV
- STOREV

The ADDV command activates an inactive mass storage volume. The volume can then be mounted by the host operating system.

The CONVERTV command copies the data from a 3336 Model 1 volume to a mass storage volume or copies the data from a mass storage volume to a 3336 Model 1 volume. This command aids in conversion of 3330 volumes to and from MSS.

The COPYV command makes a copy of an active mass storage volume. The copy maintains the same volume serial number as the original volume. The copy can either be stored in the MSF or can be ejected; in either case, it cannot be mounted by the host operating system.

The CREATEV command creates mass storage volumes from scratch data cartridges. CREATEV defines the volumes to MSS, formats the volumes for use by the operating system, and records information about the volumes in the Mass Storage Volume Inventory. CREATEV also assigns volumes to a group if requested to do so. The volumes are then accessible to the host operating system for mounting. Volume attributes such as BIND and SHARED are specified during CREATEV.

The EJECTV command physically ejects an inactive mass storage volume from the library. Both cartridges of the volume are ejected.

The MODIFYV command changes information about an active mass storage volume. The user can change one or all of the following: volume serial number, volume owner, volume description, volume group, MSS mounting attributes, use-designation, expiration date, and the number of backup copies to be retained.

The RECOVERV command restores the data from a mass storage volume copy to the original active volume or to another active mass storage volume. The copy remains as an inactive volume.

The REMOVEVR command removes one or more volume records from the Mass Storage Volume Inventory; the volumes must not be physically in the cartridge store. This command allows removal of records that were maintained in the Inventory purely for information. It also allows the removal of records for volumes that were ejected and will not be returning to the MSF.

The SCRATCHV command returns the cartridges making up a mass storage volume to the pool of scratch cartridges, and removes the volume record from the Mass Storage Volume Inventory. The volume must be either an active volume or a copy of a volume. The active volume must be empty (it must not contain any data sets) and it cannot be a VSAM candidate volume. The copy must be a copy made by the COPYV command and recorded in the Mass Storage Volume Inventory as a copy.

The STOREV command makes an active mass storage volume inactive; the volume then cannot be mounted by the host operating system. The volume can be left in the cartridge store or can be ejected for shelf storage.

Group Management Commands

Group management commands are:

- CREATEG
- MODIFYG
- SCRATCHG

The CREATEG command creates a group record in the Mass Storage Volume Inventory. SYSGROUP, a default group, is created at SYSGEN time; however, additional user-defined groups can be created via CREATEG. Following the creation of a group record in the Inventory, the user can assign to the group, via the CREATEV or MODIFYV command, those volumes to be controlled by the MSVC functions of the MSSC.

The SCRATCHG command deletes a group from the Mass Storage Volume Inventory. The group must be empty; that is, no mass storage volume, active or inactive, can still belong to the group. Any user-defined group can be scratched.

The MODIFYG command changes the group information recorded in a group record in the Mass Storage Volume Inventory. The user can change any data that can be initialized via the CREATEG command. Any user-defined group can be modified.

Report Generation Commands

Report generation commands include:

- LISTMSF
- LISTMSVI

They are used to generate status reports on the groups, volumes and, cartridges in the MSF.

The LISTMSVI command is used to generate reports on volumes and groups, using the MSVI data set. The reports show the status of the volume or group in detail, including information on available free space, owner, retention, number of free extents, etc. Options for selective reporting are:

- All volumes
- Specified volumes
- Volumes not associated with a group
- All groups
- Specified groups
- Groups that have exceeded the free space threshold

• Volumes within groups that will have expired before a specified date

The LISTMSVI command is a valuable tool for managing mass storage volumes because of the space information and retention data included in the reports. You can run LISTMSVI to determine which volumes or groups need attention. Then, using the LISTCAT command or LISTVTOC utility, you can identify problems on a data set basis.

The LISTMSF command is used to generate reports showing active volumes, inactive volumes and scratch cartridges in a particular MSF. Summary information provided includes total number of active volumes, inactive volumes, scratch cartridges, and empty cartridge locations. All report information is derived from copies of the MSC tables. The list of active volumes is in volume serial number sequence and shows each volume's mounting attributes. The list of inactive volumes is in volume serial number sequence and shows the mounting attributes of each volume plus the cartridge serial number of the first cartridge that the volume resides on. This command also shows the physical location of each cartridge in the MSF.

MSC Control Commands

The MSC control commands are:

• TRACE

• TUNE

The TRACE command sets on or off an MSC hardware trace that records activity within the MSS. The TRACE command can also be used to provide a dump of the trace activity information. This dump must be formatted by the TRACE Report utilities. The hardware traces cartridge movement, staging, and destaging.

The TUNE command displays or changes parameters that affect MSS performance. These parameters control how many inactive pages are available on staging DASD and how active pages are selected for destaging when more inactive pages are needed. The user can use this command to alter parameters which affect the MSC least-recently-used (LRU) algorithm.

For more details on Access Method Services utility commands for the MSS, see the publication *OS/VS Mass Storage System (MSS) Utilities for Space Management.*

Data Set Support Commands

The VSAM data set support commands are:

• DEFINE

• ALTER

• LISTCAT

The DEFINE command is extended by the addition of three new sets of parameters:

1. VSAM data set staging attribute

2. VSAM data set destaging attribute

3. Non-VSAM data set owner and retention period attributes

The ALTER command is also extended by the addition of the three sets of parameters described under DEFINE.

The LISTCAT command is extended to add selection criteria for listing VSAM and non-VSAM data sets. These criteria use the creation and/or expiration date of each data set and the listing obtained is for only those data sets meeting the criteria. This type of listing aids users in locating expired data sets or potentially unused data sets.

Other Utility Functions

Other utility functions include:

- DASD utilities
- Recovery utilities
- Service utilities

The primary DASD utility to support MSS is IEHDASDR. New utility control statements allow the user to specify that the pack being initialized is to be used on a staging drive as a staging pack. This affects the allocation of alternate tracks and other characteristics of the pack.

The major Recovery utilities in the MSS environment are used in the recovery of the MSVI data set. A set of utilities assists the user in the reconstruction of the MSVC data set, using the MSVI Journal data set, if reconstruction is required.

The primary Service utilities for MSS are provided as diagnostic aids to IBM Field Engineers and the MSS Space Manager. An example is the System Data Analyzer, which is described in the "Serviceability" section of this publication.

User Programs

The 3850 MSS was designed for ease of installation in the OS/VS environment. Access to application data sets is via the existing device support for 3330-1, 2, or 11 disk storage devices. Users of tape applications can look at the introduction of MSS as a device conversion to "virtual" 3330s. Tape applications that use device-independent coding techniques can be moved to the MSS with only minor JCL changes. Users of 3330 disk applications can quickly take advantage of MSS by moving data sets to mass storage volumes, and making minor JCL changes.

The primary considerations for executing user programs in the MSS environment relate to JCL, and the access method support needed by the application.

Job Control Language (JCL) Considerations

For most programs, changes to JCL DD statements are all that is necessary to convert data set definition from tape or disk to virtual 3330 volumes.

Disk JCL

For data sets presently on disk, the only change necessary in the DD statement is from "UNIT=3330" to "UNIT=3330V". If conversion is from other disk storage devices such as the 2314 or 2311, any SPACE specification in cylinders may have to be adjusted to the 3330 capacity.

A mass storage volume group designation may also be considered, depending on the user's grouping plan. The group specification is made on the DD statement, via a new parameter.

For example, the JCL DD statement change required to specify a 3330 virtual unit is:

```
from:   //PAY1      DD  DSN=A0001,DISP=(NEW,CATLG),UNIT=3330,
        //              SPACE=(CYL,(10,10))
to:     //PAY1      DD  DSN=A0001,DISP=(NEW,CATLG),UNIT=3330V,
        //              SPACE=(CYL,(10,10))
```

Tape JCL

A user's tape file JCL can vary considerably, depending on programming standards, usage of catalog options, usage of generation data groups, and external manual controls. Therefore, the changes necessary to convert the tape JCL to virtual volume JCL varies from a few parameters to many.

In the case where the data set is standard labeled and cataloged, the changes to the JCL for creating a data set are:

```
from:   //DD1 DD DSNAME=PAYMST,DISP=(,CATLG),UNIT=3400-3
to:     //DD2 DD DSNAME=PAYMST,DISP=(,CATLG),UNIT=3330V
```

The specification of a new parameter identifying a group directs the data set to a mass storage volume assigned to the group with that name. The new parameter specification results in the SPACE parameter being supplied from the group's predefined default primary and secondary space allocations. SPACE specified without the new group parameter being specified causes the data set to be allocated to a storage volume. If no storage volume exists, it will be assigned to the default-group, SYSGRP.

If tape data sets are cataloged and have standard labels, no change is necessary to the JCL statement used to retrieve an old data set such as:

```
//DD1 DD DSNAME=PAYMST,DISP=OLD
```

OS/VS Access Method Support of MSS

VSAM

Some options important in an MSS environment have been added to VSAM. Staging options allow the user to choose among three different staging procedures:

1. Staging a data set at OPEN time, which should be the normal way of operation.

2. Staging and binding a data set at OPEN time. A bound data set cannot be destaged by the MSC until the data set is unbound or the volume on which it resides is demounted. Data sets required by high performance applications should be bound.

3. No staging of data at OPEN time. Data is staged one cylinder at a time as required by the host CPU(s). This is called "cylinder fault" mode and could be used for infrequent accesses to very large data sets.

Destaging options permit the user to specify two procedures:

1. Synchronous destaging of a data set at CLOSE time. This option causes the modified cylinders to be completely destaged before control is returned to the user program. If hardware errors are encountered during the destaging process their presence is communicated to OS/VS allowing immediate recovery action.

2. No destaging of a data set at CLOSE time. Modified cylinders are destaged by the least-recently-used (LRU) algorithm of the MSC or when a volume is demounted.

These options are recorded in the VSAM catalog so that they can be used without program or JCL change.

ISAM

No modification has been made to ISAM. This means that there is no support for the staging of ISAM data sets by OPEN. All data is accessed in "cylinder fault" mode. ISAM data sets should be converted to VSAM data sets and processed using the ISAM-VSAM compatibility function of VSAM, which eliminates accessing data in cylinder fault mode.

Other Access Methods

BSAM, QSAM, BDAM, BPAM, EXCP are changed to support staging at OPEN time. But the "BIND" and "CYLINDERFAULT" options are not supported for these non-VSAM data sets. These data sets are always staged at OPEN time.

Figure 13 shows the operation of the common access methods.

	Operation	VSAM	SAM, DAM, PAM, EXCP	ISAM
S T A G E	**Stage** Data Set at OPEN Time	Yes (optional)	Yes	-
	Stage and Bind Data Set at OPEN Time	Yes (optional)	-	-
	Cylinder Fault Access to Data Set (No Staging at OPEN)	Yes (optional)	-	Yes
D E S T A G E	**Synchronous Destage** of Modified/New Data Cylinders at CLOSE Time	Yes (optional)	-	-
	Asynchronous Destage of modified/new data cylinders based on Least Recently Used (LRU) or volume de- mount	Yes (optional)	Yes	Yes

Figure 13. OS/VS Access Method Stage and Destage Options

System Availability and Data Integrity

Computer-based systems have become increasingly important to businesses, governments and other organizations. For this reason, increasing attention has been given to systems availability and data integrity.

System availability is having the system when you want it. Elements of system availability are:

- *Reliability* of the components.
- *Continuity* of system functions
- *Serviceability* of the components

Reliability means that the system stays up. When components become unreliable, they may disrupt continuity of system functions. When that happens, the serviceability of the components becomes an important consideration.

Reliability is the measure of the probability—at best an estimate—that the system will do what it is designed to do for a given period of time.

Continuity involves answers to:

- How often will a malfunction occur?
- What effect will malfunctions have on system capability?
- How long will it take to fix?
- How often will preventive maintenance be performed?
- When will engineering improvements be installed?

Serviceability questions are phrased as how easy and fast is a system to fix and maintain and what effect does restoring one element have on another.

Data Integrity means the ability of the system to store, maintain, update, and move data without alteration due to a malfunction. Important data integrity considerations are:

- *Recovery* of data after system failures.
- *Data Security* from unauthorized access, change, or exposure.

Recovery is an automated function in the sense of returning to a normal state after a temporary error in the system. Recovery is also the process of manually restoring a system when it is unable to recover automatically.

Data security refers to protecting data from unauthorized disclosure, modification or destruction by accidental or intentional means.

System Perspective

To look at these aspects of a specific computer-based system, both IBM and the user must have the same understanding and perception of the system. A *system* is not merely an interconnection of computer hardware units, software, and firmware. A system should be perceived—by everyone concerned with it—as a dynamic combination of resources:

- Hardware
- Software
- Firmware
- Data
- Operators
- Application programs
- Normal operating procedures
- Contingency procedures
- Computer room environment
- Management

Note that a key concept is that all systems continually undergo change. A system running a payroll at ten o'clock, for example, is different from the "same" system running linear programs at eleven o'clock. Even the hardware changes, because what was a critical unit for the payroll (for example, a printer) is possibly not even used in the linear-programming application.

When a question is asked regarding the systems availability or integrity characteristics of a system, the questioner is really asking "management" questions about the control of a full set of resources to do his job. The key to meeting security, integrity, and system availability requirements is not only good hardware and systems design, but also includes good availability management, contingency procedures, well trained operators, and intelligent planning.

IBM's 3850 Mass Storage System, properly employed with good systems design, provides an opportunity for improved resource management over conventional tape/DASD in these areas:

1. Tape handling is completely automated. This reduces problems caused by manual handling as dirty leaders, damaged tape, bent reel flanges, etc.

2. Improvements in the recording quality of the tape, reduction in the recording head-to-tape separation, and advances in error correction code provide not only. improved data density on the cartridge but more reliable reproduction of the data compared to tape/DASD.

3. Because much more data can be under system(s) control when compared to tape/DASD, normal operating occurrences such as misfiled, mislabeled or lost reels of tape and mismounts are nearly eliminated.

4. Contention for a unique tape data set is also reduced because the data, once staged, can be shared among two, three or four operating systems from DASD, assuming that the data set has been qualified as sharable.

5. Extensive hardware replication and sophisticated, automatic error recovery procedures are integral to the design of the Mass Storage System.

6. The Mass Storage System has been designed so that most maintenance can be accomplished concurrently with productive use of the system.

Mass Storage System Availability

Description of MSS Availability must start with a review of the general architecture of the system. The following schematic (Figure 14) shows a Mass Storage System consisting of one 3851 Mass Storage Facility. The MSS is composed of

eight major groups of components indicated by the columns of boxes. The numbers in the schematic indicate the maximum number of each type of component that can be included in an MSS containing one Mass Storage Facility.

Internal to 3851 MSF Frame   DASD

Mass Storage Controller No. 1	Right Accessor	Data Recording Controller and DC Power	Pneumatic Supply and AC Sequencer	Data Recording Devices (DRD)	3830-3	3333	3330
1 of 2	2 of 2	up to 4	up to 4	up to 8	up to 8	up to 32	up to 128 (staging drives)
Mass Storage Controller No. 2	Left Accessor	Data Recording Controller and DC Power	Pneumatic Supply and AC Sequencer	Data Recording Devices (DRD)	3830-3	3333	3330

Figure 14. IBM 3850 Mass Storage System

The MSS is extensively replicated. Any cartridge can be mounted on any data recording device which is accessible to the 3830-3 selected by the MSC. Therefore, with alternate paths in a system configured for availability, the impact of any single unit failure and many multiple unit failures will not cause the system to stop. Rather, it will operate at a degraded data rate until the failure is corrected.

Mass Storage Facility—Redundancy in Data Paths

Two accessors and their associated controls and power supplies are included in each MSF. If one fails, an error recovery program in microcode in the MSF causes the functioning accessor to push the failing accessor into a "garage" to await service. The remaining accessor continues to operate, and the effect in Models 2, 3, and 4 is to slightly reduce cartridge access rate. The failing accessor can normally be serviced in its "garage".

Except in Model 1, each data recording device is cabled to its primary DRC and to another DRC (Figure 15). Each pair of data recording devices is cabled to its own power supply and pneumatic system. If power or pneumatics fail, only that pair of DRDs is disabled. In the models 2, 3 and 4, the system continues to operate with a potentially degraded data rate.
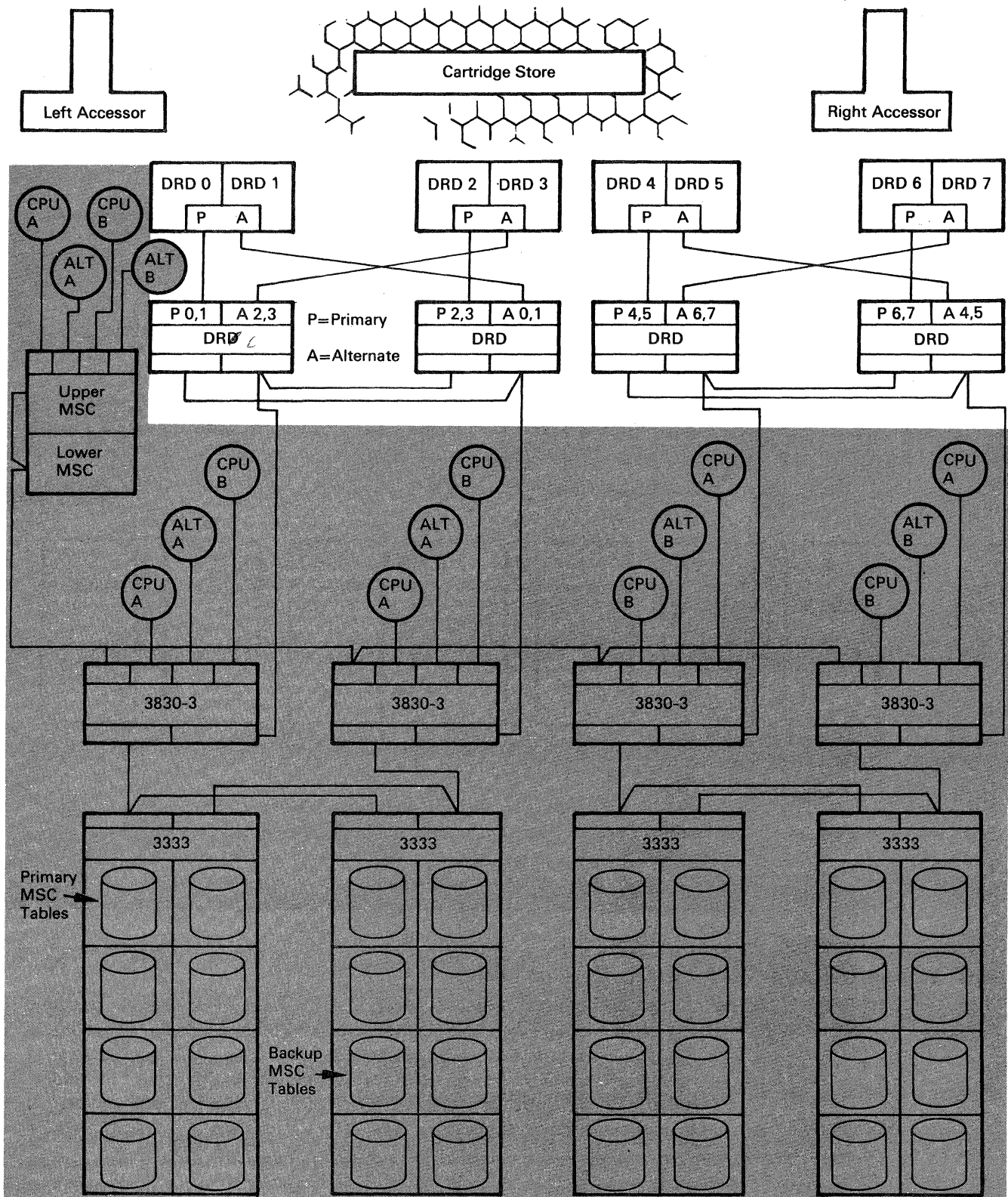
Notes:

1. Each 3830-3 is accessible to two Data Recording Controls.
2. One host system can be down and the other is operational.
3. The configuration can lose any one 3830-3 and still be operational.
4. The configuration can lose any one 3333 and still be operational.

Figure 15. Mass Storage Facility Data Path Redundancy

If the primary DRC fails, the switch to the backup DRC is automatic.

In turn, each data recording control can have access to two 3830 Model 3 Storage Controls and each Storage Control can attach to up to four DRCs. DRCs also have independent power supplies, so that if one power supply fails it does not disable the other DRCs in an MSF.

For each of these situations, reconfiguration is automatic through error recovery procedures in the MSC microcode without dependency on the host system(s). The failing unit is marked unusable and the operator is notified. He may vary it offline. The Mass Storage System is designed to permit maintenance and check-out on the unit varied off line without interrupting the rest of the storage system.

Mass Storage Facility—Redundancy in Control Path

Each Mass Storage Facility contains one Mass Storage Control (MSC) and can optionally contain a second one (Figure 16). In the event of an MSC failure, switchover to the backup MSC is automatic. Only one is active at a time. Both MSCs are powered and memory is loaded at power-on time of the MSF. The MSC which has the lower interface becomes the primary and completes an initial microprogram load (IML). The backup MSC loops in a master dispatcher loop until it receives one of a select group of commands.

When an MSC failure is detected by a host, operating system error recovery procedures issue commands to switch from the on-line MSC to the backup MSC. The backup MSC then completes the same IML as the first MSC. Upon completion of the IML, a Device End is issued. The operating system will then issue the INITIALIZE and HOST READY FOR MESSAGES orders. Upon completion of these orders, the CCW that failed is retried, now using the backup MSC.

This entire switchover procedure occurs in seconds and will not interrupt jobs already in process on the host CPUs. Control messages that were in process at the time of failure of the first MSC will be reestablished by the backup.

This switchover procedure operates identically between the two MSCs in a "B" series Mass Storage Facility or between each of the MSCs in two "A" series Mass Storage Facilities.

Customer Engineering action is necessary to check out the failing MSC, take appropriate repair action, and again load its memory. This is done without interrupting the use of the Mass Storage System.

DASD

When configured for availability, the DASD allows for alternate data paths between each Data Recording Control and 3830-3 and from there to each 3333 Disk Storage and Control. Failure to properly configure may cause an unnecessary reduction in available data paths, thereby reducing the storage system's capacity to deliver data when operating in a degraded mode or even disabling the MSS if the Mass Storage Control tables are not accessible.

Each 3333 should be cabled to a pair of 3830 model 3 Storage Controls and each 3830-3 should be cabled to a pair of data recording controls.

Figure 17 shows an example of a complex configuration involving two hosts, four 3830-3's, four 3333's and 32 disk drives.
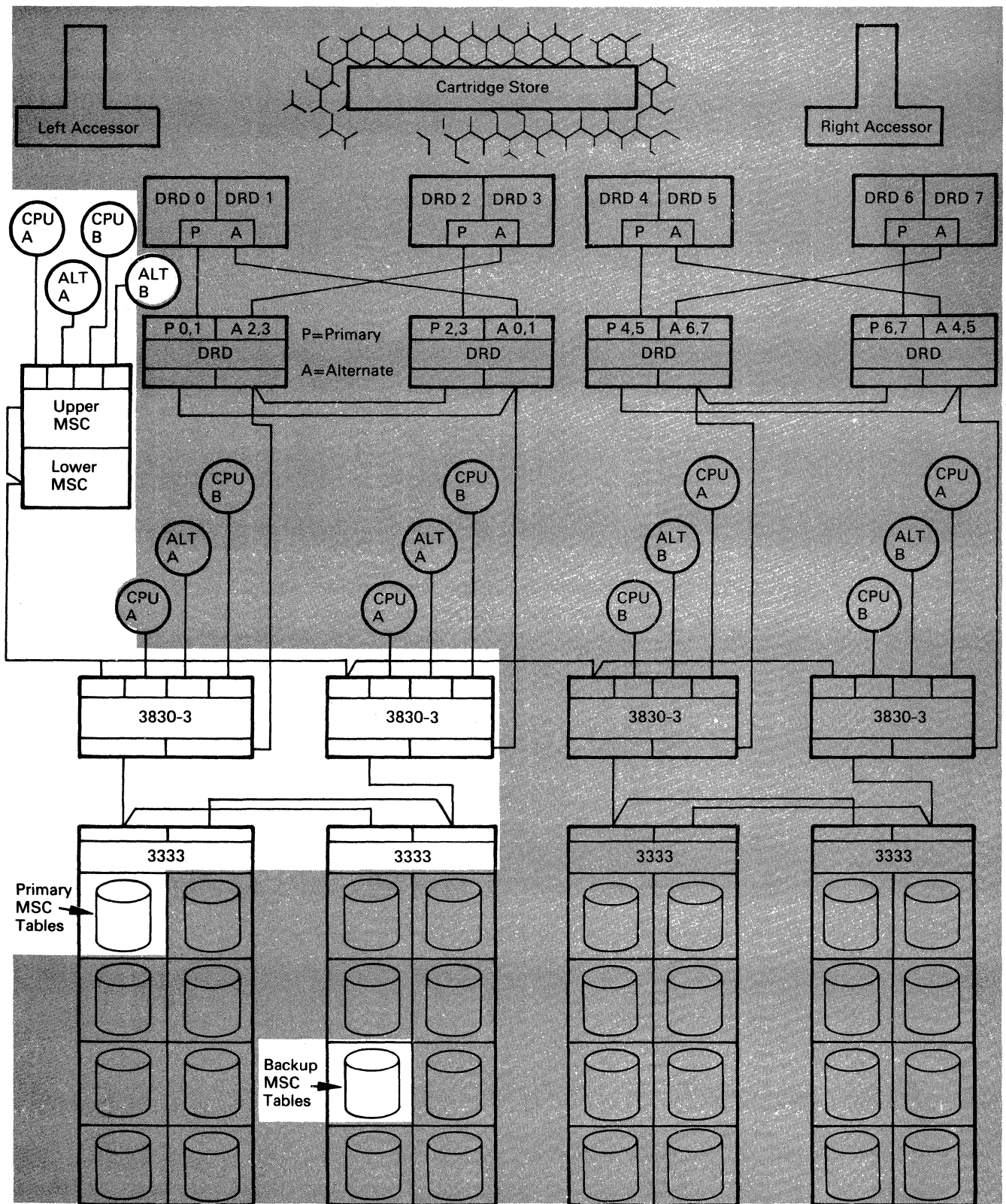
Notes:

1. Each 3830-3 is accessible to two Data Recording Controls.
2. One host system can be down and the other is operational.
3. The configuration can lose any one 3830-3 and still be operational.
4. The configuration can lose any one 3333 and still be operational.

Figure 16. Mass Storage Facility Control Path Redundancy
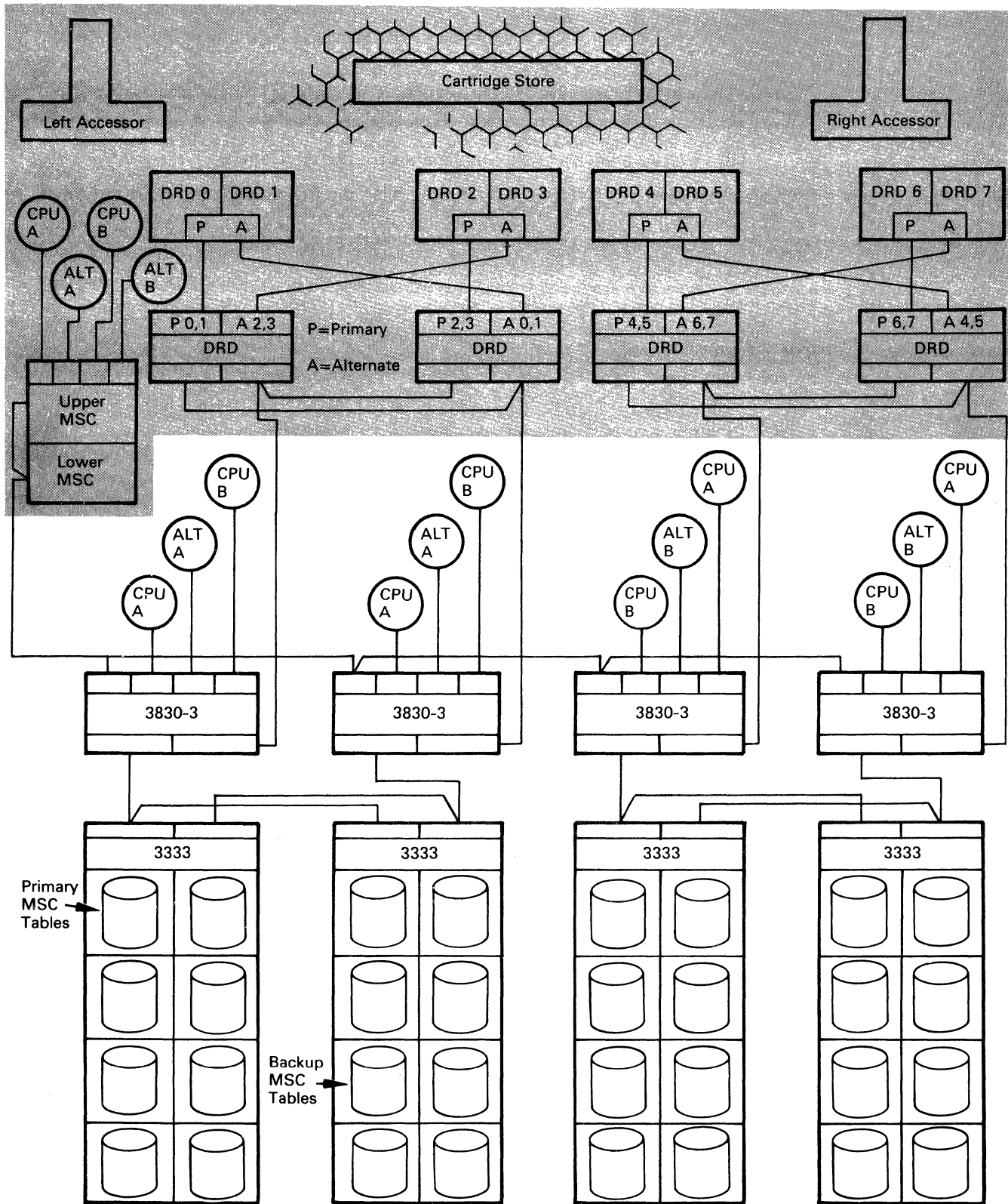
Notes:
1. Each 3830-3 is accessible to two Data Recording Controls.
2. One host system can be down and the other is operational.
3. The configuration can lose any one 3830-3 and still be operational.
4. The configuration can lose any one 3333 and still be operational.

Figure 17. Complex Mass Storage System DASD Configuration

Unduplexed Configurations

Note that the following configurations do not provide maximum availability because they include unduplexed components.

1. A-series models contain only one Mass Storage Control.

2. Models A-1 and B-1 contain only one: data recording control, DC power supply, AC sequencer, and data recording drive pneumatic supply for the Mass Storage Facility.

3. Single 3830 Model 3 storage control subsystems do not provide an alternate path around the 3830.

Mass Storage Control—Table Access

The MSC expects to find its tables at one of two fixed locations. The primary table pack location is 3830 (address 0), 3333 (address 0) drive (address 0). If the primary location is not available to the Mass Storage Control, a duplicate set of tables exists on a disk drive attached to a separate 3333. During operation, the MSC tables at both locations are maintained so that they are always exact copies of each other. See Figure 18.
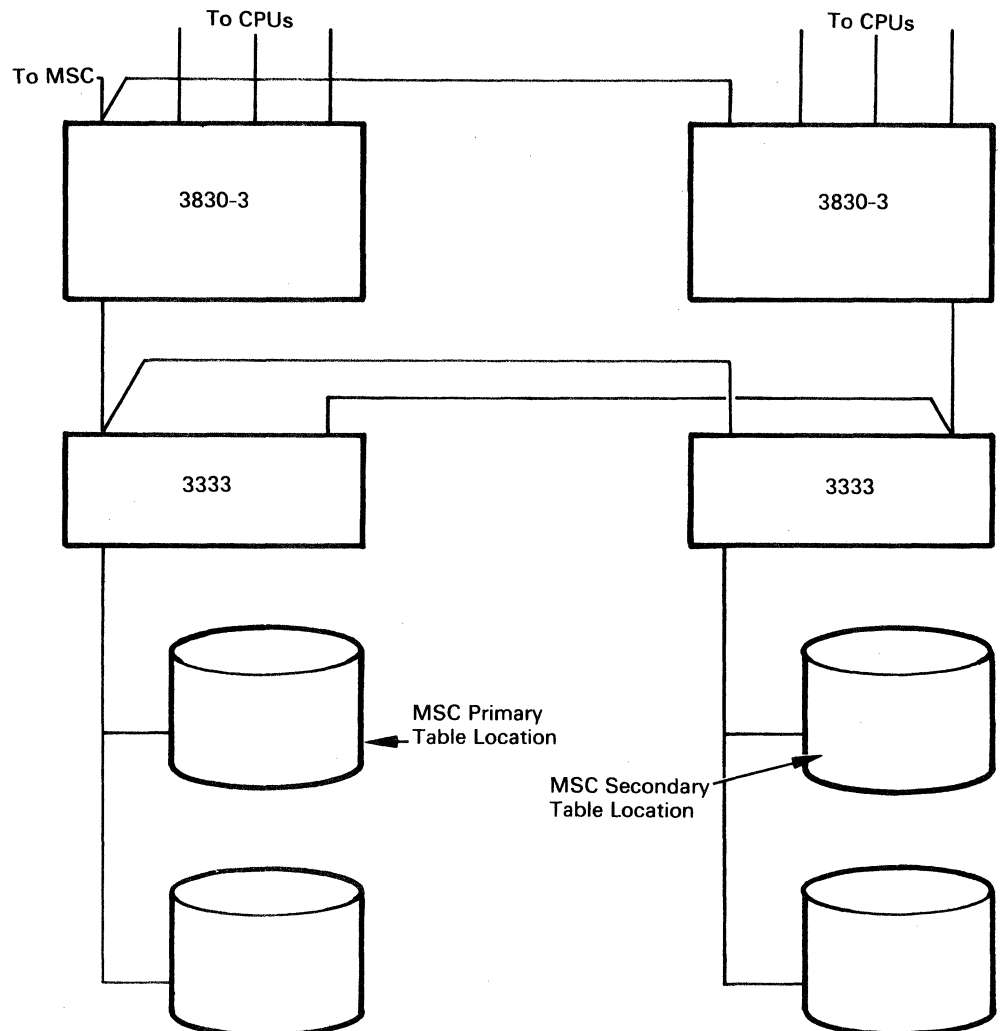
Figure 18. MSC Duplicate Table Pack Locations

Disk Pack Swapping

A disk pack can be removed from one staging drive and placed on another of the same model, as is possible with current DASD, to assist in preventive maintenance and help resolve intermittent drive failures. To accomplish this, you must turn off the drive containing the disk pack to be swapped and vary off another staging drive or a designated spare drive on the same 3333 string. The vary off command causes the data on a staging drive to be destaged. Data in use will be restaged on another staging drive. Move the address plug, the pack and VARY ON the new drive.

Serviceability

Serviceability is important because it is a positive influence on availability. The better the design for serviceability of a system, the smaller the impact on availability of failures within that system. A key design objective for the 3850 Mass Storage System has been to maximize concurrent maintenance. In fact, where a Mass Storage System configuration provides replicate components, almost all scheduled and unscheduled maintenance, including problem diagnosis and checkout, can be done while the storage system is in use. This overlap is achieved by varying off only those components required for service. The system continues to operate in degraded performance mode. Another design objective has been to provide better maintenance tools to reduce "long calls". These objectives are served by:

1. Comprehensive Maintenance Library Manuals that are effectively cross referenced and easy for Customer Engineers to use.

2. Improved error checking in the hardware, when compared to previous IBM products, and a more extensive display of sense data.

3. New microdiagnostic programs which are entered in the MSC and 3830 Model 3 by the Customer Engineer from a diskette using a CE console. This is done while the MSS is operating in a normal environment.

4. Improvements to On Line Tests (OLTs) so that the diagnostic intelligence is concentrated in microdiagnostic programs below the host CPU. Some of the functions provided by the improved diagnostic programs are to:

   - Isolate different types of failures

   - Report cartridge store addresses and conditions of access errors

   - Measure accessor motion

   - Report addresses and the number of control checks

   - Analyze paths

   - Test interfaces

   - Update microcode

5. A program called System Data Analyzer (SDA) designed to help the Customer Engineer diagnose existing problems more quickly and anticipate failures. Sense data, buffered-usage-log data, and data logged in exceptional circumstances are automatically entered into the SYS1.LOGREC data set during MSS operation. Because (1) this data is voluminous, and (2) the interrelationships of MSS components are complex, the program is valuable in processing logged information.

Although the familiar EREP reports include MSS information, SDA is the primary vehicle for conveying MSS logged data because the SDA reports are concise, easy to use, and informative.

SDA reports are aimed at two audiences:

- The user who needs the Data Cartridge Statistics report to manage his MSF cartridge inventory.

- The Customer Engineer who uses all SDA reports to stay aware of hard and soft errors, incidence of certain exceptional conditions, apparent trouble spots, etc., during his planned and unplanned MSS service activity.

The 14 SDA reports fall into three categories:

1. Volume Error Statistics.

   The only report in this category is Data Cartridge Statistics, which provides a count of temporary read and write errors versus total bytes passed for each by cartridge.

2. Analytical interpretation, in which SDA does considerable detailed analysis to report conclusions about real or potential problems to the CE.

   Reports in this category are:

   - Path Analysis

   - Fault Symptom Code Summary

   - SDA Input Summary

3. Data Reduction, in which SDA collects and presents related information so as to minimize the time the CE needs to understand it.

   Reports in this category are:

   - Data Handling Usage for DRDs

   - Data Handling Errors

   - Cartridge Store Buffered Log

   - Cartridge Store Forced Log

   - Equipment Checks

   - Power Supply

   - Device Checks by DRD

   - DRD Index Program and Thread Program Counter Summary

   - Summary of DRC Equipment Checks and Data Checks

   - Summary of Data Handling Checks by DRC and DRD

SDA is an SCP problem program that runs under VS1 or VS2. Input to it is SYS1.LOGREC, either directly or via EREP accumulation (history) data sets. In the latter case, SDA need not be run on a system to which the MSS is attached. Where appropriate, SDA's reports direct the CE into the MSS Maintenance Library Manuals for further diagnosis or repair procedures.

The improvements in serviceability result from hardware designed for minimum maintenance and for faster problem determination or anticipation and from an enhanced set of diagnostic tools. The value of improved serviceability is that it enables IBM to reduce the maintenance impact on data processing operations.

Data Integrity

The integrity of data stored in the 3850 Mass Storage System is maintained by:

- The extensive use of checking circuitry and diagnostics.
- Parity.
- A new error correction method, Extended Group Coded Recording, for recording data on cartridges.
- DASD error correction coding.
- Multiple recording of cartridge serial numbers.
- Automatic reconstruction of 3830 Model 3 tables.
- Maintaining duplicate Mass Storage Control tables.

Checking and Diagnostics

Data circuits in the data recording control, data recording device, 3830 Storage Controls, Integrated Storage Controls, 3333 Disk Storage and Control, and the 3330 Disk Storage Drive are hardware checked while in use. When they're not in use, a Loop Read to Write diagnostic command checks the data recording control circuits in the Mass Storage Facility. A similar diagnostic command checks data circuits in the DASD subsystem.

Physical travel of each accessor is monitored and the actual physical location address is compared with the address in the command sent to the accessor control that initiated accessor movement. This ensures that the accessor is selecting from the intended location. Each accessor is also equipped with a series of physical interlocks to prevent damage in the event of a failure in the accessor control system.

Parity

Odd parity is maintained in the data flow through the Mass Storage Facility.

Extended Group Coded Recording

Data is recorded on cartridges using the Extended Group Coded Recording method. This error correcting code and associated circuitry can correct up to 32 of 208 bytes on-the-fly.

DASD Error Correction Code

As data is transferred from the host channel or from the MSF to disk storage (write operation), the 3830 Model 3 storage control removes the parity bit associated with each byte. It then computes the error correction code bytes, which are written after each recorded area. The correction code bytes, coded to represent the data in the recorded area, are used for both error detection and correction.

As data is transferred from disk storage to the channel or to the MSF (read operation), each area is inspected by the 3830 Model 3 storage control and the error correction code bytes are recalculated for each area. The 3830 correction code corrects single bursts of 11 bits or less.

If a correctable data error is detected in the home address, count, or key areas, the 3830 Model storage control internally executes the error correction function through through the use of command retry. If an error in the data area is detected, the host correction function is determined by the system error recovery procedures.

The correction code bytes are removed and proper parity is generated by the 3830 Model 3 storage control before the data is transferred to the channel or to the MSF.

Multiple Recording of Cartridge Labels

When tape is threaded on a data recording device, the cartridge label is read and compared by the Mass Storage Control with the required cartridge label contents. An unequal compare automatically invokes microprogrammed error recovery procedures; an equal compare allows the cartridge to be used. Because of the importance of this label, it is recorded twice in the cartridge label area. The cartridge serial number is also imprinted on the tape so that it can be visually inspected through the transparent cartridge shell.

3830 Model 3 Tables

The 3830 Model 3 contains tables in a control store area which serve the following purposes:

- The Virtual Address table provides information about each virtual address assigned to the 3830-3.

- The Virtual Volume Information table tells about each Virtual Volume mounted such things as whether it is read only, busy, shared and/or reserved.

- The Page Status table allows the conversion of virtual page numbers to real page numbers and virtual unit addresses to logical unit addresses.

- The Logical to Real Conversion table allows the conversion of logical unit addresses to real unit addresses.

Although these tables reside in the 3830-3, they are maintained by the Mass Storage Control. Should a malfunction occur in the 3830-3, the Mass Storage Control refreshes the tables without manual intervention.

Duplicate Mass Storage Control Tables

Duplicate MSC tables are maintained in a reserved space (32 cylinders) on each of two staging drives. These tables are:

Table	Function
Volume Inventory	Records cartridges in mass storage volumes in the Mass Storage Facility. It also describes a volume identifier, its attributes and its physical location.
Scratch Cartridge	Lists the cartridge serial numbers and their locations which do not have a VOLID assigned.
Mounted Volumes	Shows virtual volume mounted in response to host commands.
Staging Drive Groups	Shows the relationship of staging page and unit addresses to virtual pages and volumes.
Transient Volumes	Shows volume identifiers for inactive volumes.
Configuration Map	Shows the relationship of CPU(s), channels, MSF and DASD.
Virtual Volume Address—VOLID Cross Reference	(Identification Cross Reference)—relates a virtual address to a volume identifier "assigned to" address.
Trace Tables	Time stamps the start of a TRACE command and contains the two trace areas.
Verification Tables	Used during system verification. Contains a description of all table locations in addition to control blocks which describe the available system configuration.
Schedule Queues	Shows the schedule queues for the MSF and for DASD.

In addition, a recovery journal is maintained where all control record(s) to be updated are written in their original form. If an order completes correctly, the journal is nulled. If the order is not completed successfully, the above tables are restored to their original content by reapplying records in this journal. In the event of a Mass Storage Control failure, the next IML (or the switch-over procedure between MSCs) will restore these tables.

The Mass Storage Control must have access to its tables to operate. These tables, including the recovery journal, are maintained in duplicate on two staging drives in the same staging drive group, as shown in Figure 18.

Note that if any one of the 3830s or 3330s or staging drives in the data path to the MSC tables fails, there is an alternate path to the tables.

There are two potential problems with regard to MSC table integrity:

Partial updates   Caused by errors in processing a given command or by MSC failures.

Media failures   Caused by DASD I/O errors.

In the case of a partial update the MSC does not receive a "completion" to its order. When this happens, the contents of the recovery journal are applied to those tables affected and the order is retried by the CPU. In the case of a media failure on a staging drive containing the primary MSC tables an identical record is available at the secondary location. If a disk pack containing the MSC tables has been damaged, the tables from the remaining good pack should be copied immediately.

Recovery

There are many unique error recovery routines implemented in microcode in the Mass Storage System. Some are completely transparent to the user and some require operator intervention. The list in the following table is representative:

Error	Recovery Action
Read error from cartridge	If the error persists through error correction code, the read is retried. If the error persists, the tape is unthreaded and rethreaded. If the error continues to persist, the cartridge is reloaded on another Data Recording Device.
Write error to cartridge	The write is retried a prescribed number of times. If the stripe doesn't record correctly (permanent write error), the stripe is demarked and an alternate stripe recorded on. There are at least four alternate stripes for each cylinder on a cartridge. If no alternate stripes are available, then the contents of the cartridge are copied to a scratch cartridge using the Access Methods Service Cartridge Copy Utility and the failing cartridge is ejected from the MSF.
Cartridge load or thread failure	The load or thread is attempted again on the same Data Recording Device. If the error persists, the load-thread cycle is attempted on an alternate Data Recording Device.
Accessor failure	The failing accessor is automatically "pushed" into its garage by the operational accessor. The console operator is notified and in Models 2, 3, and 4 the system operates in a potentially degraded mode. Model 1 operates without degradation.
Data Recording Control failure	After a prescribed number of unsuccessful retries the DRC is automatically marked unusable and the operator is notified. The system operates in a potentially degraded mode.
Data Recording Device failure	After a preset number of unsuccessful retries, the DRD is marked unusable automatically and the operator is notified. The system operates in a potentially degraded mode.
3830-3 failure	The operation is retried and if the error persists, the operator is notified.
Mass Storage Control Failure	Automatic switchover triggered by OS/VS to a backup MSC (if "B" series or dual "A" series). Control messages in process are restarted.

Recovery When a Host Fails

If a host system fails, there are certain activities required of an operator that are unique to the Mass Storage System. If, in a multi-host environment, the failing host has been designated the primary host, the operator must designate a new primary host in order to permit MSC/Host messages.

When a host that had been communicating with MSS fails, it is highly probable that some MSS resources had been allocated to the failing system. Before these resources can be reallocated, the operator must issue a PURGE command. This will cancel all outstanding requests and release any resources allocated to the failing host.

Backup

Good systems design requires backup data both for data integrity and to provide a resource for recovery when the MSS is unable to recover automatically or when media damage occurs.

The MSS and host operating systems provide:

- Mass Storage Volume backup
- Mass Storage Volume archiving
- Generation Data groups
- Data set copy

Mass Storage Volume Backup

Volumes within the Mass Storage Facility can be classified as active or inactive. Active volumes can be mounted by the host operating system(s); they appear on the volume inventory list. Inactive volumes appear on the transient volume list and cannot be mounted by the host operating systems. The volume management commands provide an optional facility for semi-automatic backup of mass storage volumes. The number of backup copies to be retained can be specified for a volume when it is created (CREATEV command). The backup number can subsequently be changed (MODIFYV command). Thereafter, the COPYV command automatically retains the specified number of copies. Each copy made by the COPYV command is flagged as a backup copy if backup is specified. If the specified number of copies do not exist, the COPYV command creates one new backup copy. If the specified number already exists, the COPYV command automatically replaces the oldest backup copy with a new copy. If a volume needs to be restored from a backup copy (made active), the space manager can either specifically identify the copy or call for the latest backup copy (RECOVERV command).

Mass Storage Volume Archiving

Volume management commands, together with MSVC, provide assistance in archiving data sets and volumes. Within any mass storage volume group, the space manager can define one or more volumes as restricted volumes (CREATEV command).

When a space manager wishes to archive a data set, he can use existing utilities to move it to the restricted archive volume and recatalog it. When an archive volume gets full, the space manager can make that archive volume inaccessible and optionally remove it from the Mass Storage Facility for shelf storage.

Generation Data Groups

At catalog create time within OS/VS1 (OS Catalog) and OS/VS2 (VSAM Catalog and OS Control Volume) you can specify the number of generations of a data set to be retained.

Data Set and Volume Copy

Data sets can be copied using existing utility programs. The copied data can then be destaged to a different mass storage volume group in the MSF.

Volume backup and volume archiving do not require host CPU resources or channels except to update the Mass Storage Volume Inventory (MSVI) data set. The maintenance of generation data groups or the use of data set copy utilities requires system resources.

In some cases, it may be desirable to back up mass storage volumes to tape because of extremely long archival retention periods or to facilitate interchange to other systems. Also, tape, especially when recorded at 6250 bytes per inch, is more economical for very large data set storage.

Data Security

Data security is a technical question to be resolved by the use of management-approved and enforced system security procedures, manual procedures, and attention to the physical environment. Data security is an environmental consideration which involves the system functions of:

Identification   Provides the system with a unique, recognizable name for each person, device or system resource attempting to access a system resource.

Authorization   Controls the access of each person, device or system component to a system resource.

Audit   Provides a record of all accesses and attempted accesses to a system resource for review, enforcement and evaluation of approved security measures.

System Integrity   Is the ability of the system to protect itself from unauthorized attempts to compromise or bypass security controls or features.

The IBM 3850 Mass Storage System and OS/VS support these systems functions through password authorization of utility commands, by providing the mass storage volume attributes DASDERASE and READONLY, by providing RPQs designed for physical protection, and by placing more data under system control.

Password Protection

All the utility commands are authorized, because the Mass Storage System Communicator requires an APF Authorization. A user should generally place the utilities in an APF authorized, password-protected library.

The Access Method Services utility commands check data set and VSAM catalog passwords when performing volume operations. For each non-VSAM data set on a volume, the utility commands require the *read-level* password when copying or converting the volume and the *write-level* password when modifying the volume serial number, making the volume inactive, or replacing the volume with a backup copy. For each VSAM data set on a volume, the utility commands require the *master-level* password when copying or converting the volume, making the volume inactive, or replacing the volume with a backup copy. In the above cases, you can specify the master-level password of the VSAM catalog that owns the volume and the utility commands will bypass data set password checking for the VSAM data sets.

Data set passwords are not checked when a copy of a volume is scratched, because the copies cannot be mounted. The SCRATCHV commands, however, can ensure that the volume being scratched is a copy from the volume records maintained in the Mass Storage Volume Inventory.

DASDERASE/READONLY

The DASDERASE Volume attribute causes the Mass Storage Control to write over with zeros all cylinders of that volume on DASD when it is destaged or demounted. The Data Recording Control always writes over the remaining unused stripes of a cartridge cylinder allocation when destage takes place.

The READONLY attribute causes a unit check when a write attempt is made to a volume.

Physical Protection

RPQs are available to detect fire and fumes and to trigger a user-supplied fire suppression system. Also, RPQs are available to provide locks on all access panels in the Mass Storage Facility (requires the fire/fume detector RPQ) and to detect a foreign magnetic device introduced into the system.

Physical security of data is enhanced by placing the data continuously under system control. Note the contrast in Figure 19.

MASS STORAGE FACILITY
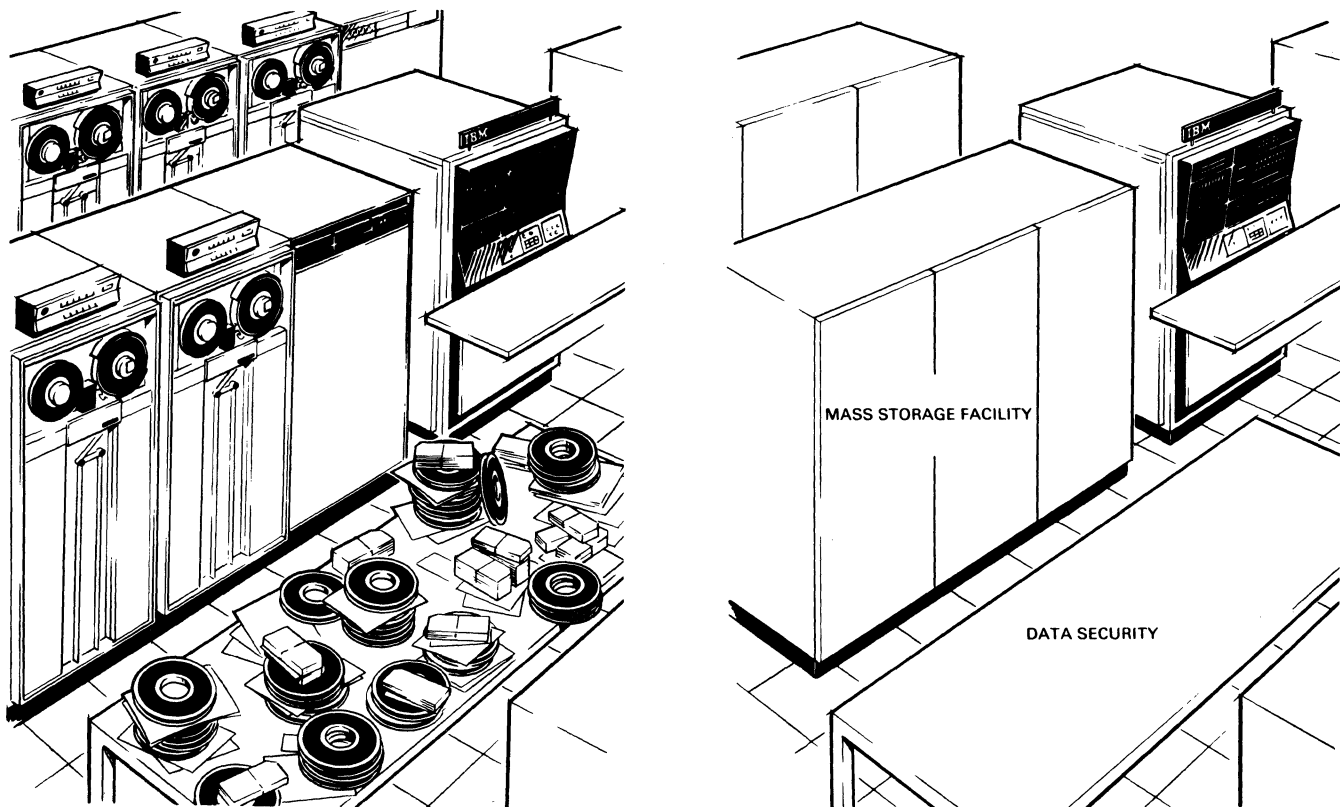
DATA SECURITY

Figure 19. Physical Data Security.
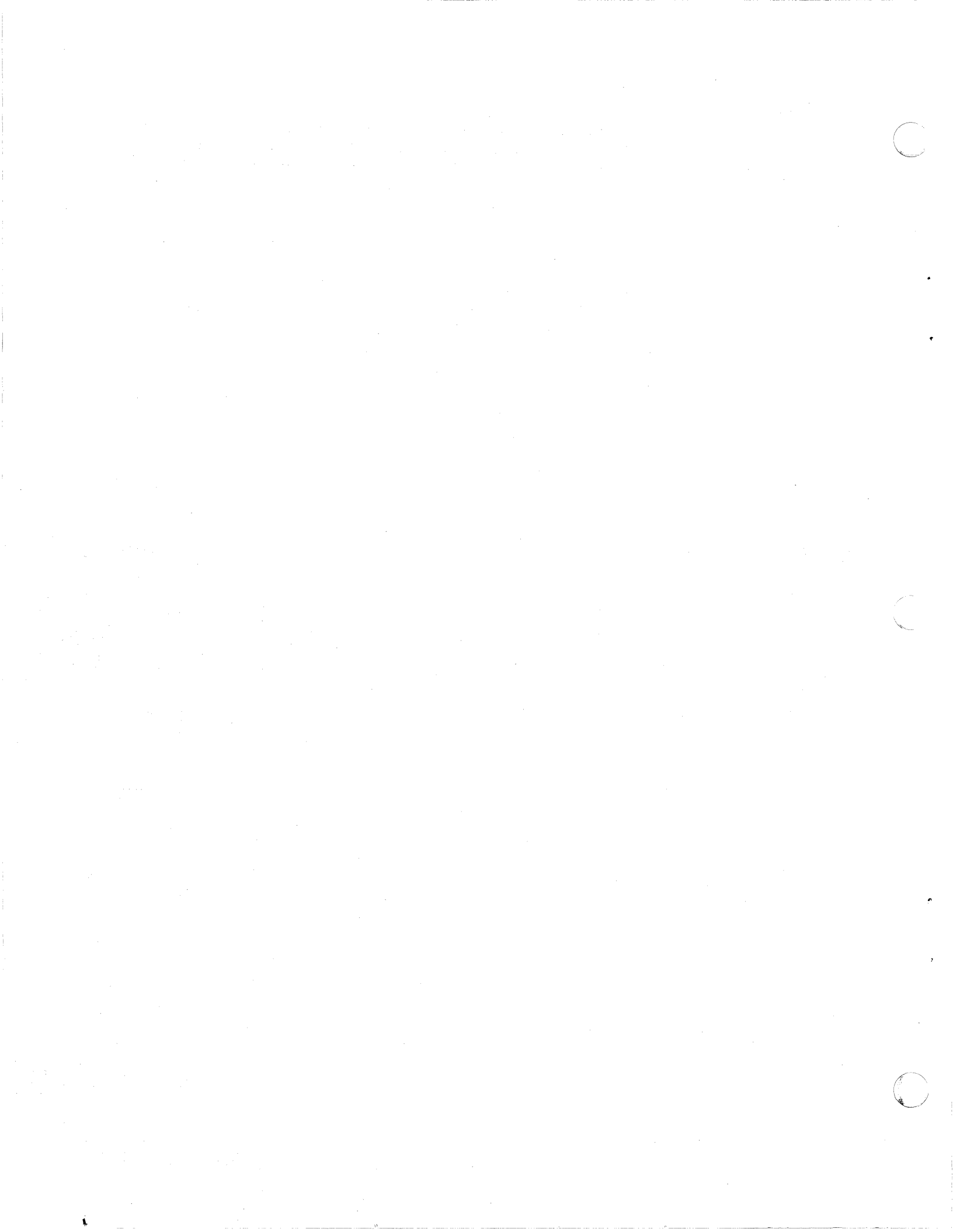
Conversion

The 3850 Mass Storage System is an extension of the virtual storage concept to data storage. Potential users of the MSS are installing central processors with virtual storage and VS1 or VS2 now. They are also engaged in expanding their application bases.

Conversion of existing applications to the MSS can usually be viewed as an I/O change, thereby conserving resources for continued application expansion. The initial conversion should appear as a tape to DASD conversion, one that is generally familiar.

The design objectives for MSS to ease conversion are:

- Use current production source code without redesign or reprogramming.

- Use current programs without re-compiling or re-linkediting.

- Limit effort to minimal Job Control Language (JCL) changes.

MSS for most users will meet these objectives. This is especially true for users who maintain and develop installation standards.

Compatibility Considerations

The compatibility objective is to provide a means by which current applications, using tape or 3330 media, can be stored by the Mass Storage System with minimum effort and expense. From a conversion viewpoint, this has been achieved by making the MSS compatible with the 3330 Disk Storage. With the exception of device-dependent code, the conversion has been reduced to converting a tape data set to a disk data set.

The virtual volume concept is compatible with present 3330 DASD operation. The virtual volume concept is simply an expanded addressing system that permits the host CPU(s) to address and access mass storage volumes as if they were additional 3330 volumes, automatically mounted on demand. The conversion to the Mass Storage System is essentially a conversion to 3330 DASD. In conversion, the user does not have to concern himself with managing the data below the 3330 (data flow between the MSF and the 3330).

For the application programmer, the Mass Storage System is an IBM 3330 with nearly unlimited space. For the systems programmer, the expanded addressing system does not change the parameters he must review; it only increases the total number. Conversion planning is affected because the volume attributes of BIND, EXCLUSIVE, and DASDERASE are new considerations which must be analyzed.

Because the Mass Storage System is compatible with the IBM 3330, the user can convert and test his tape media program/data prior to installation.

After the MSS has been installed, the CONVERTV utility destages the converted tape data from the 3330 disk(s) in 3330 Model 1 or 2 format to the MSS with a minimal involvement by the host system and no effort by the programmer. He need only make JCL unit changes from 3330 to 3330V prior to the next execution of his program. If he has used a generic unit name, his effort is even less.

Once the CONVERTV utility operation is initiated by the host system, the Mass Storage System controls the destage data flow. During this special conversion operation (3330 to 3330V), the 3330 drives are offline to the host CPU and to the Mass Storage Control. The conversion mode of operation can also be used to stage data to 3330 drives from 3330V volumes. This unique characteristic is especially useful because data can be converted during normal system operations. During conversion, it can be used for backup purposes.

VSAM/ISAM Relationships

The access method and catalog recommended for use with Mass Storage System is the Virtual Storage Access Method (VSAM). The current OS/VS access methods also function with the MSS.

Many users will install VSAM and convert at least their Indexed Sequential Access Method (ISAM) data sets. With Access Method Services Utilities, sequential and indexed sequential data sets can be easily converted to the VSAM format. The compatibility routines for existing customer programs, using ISAM, provide access to the VSAM data sets.

ISAM can function with the MSS, but only in a cylinder fault mode. The Mass Storage System will stage a cylinder of data on a cylinder fault. This option can be used for low volume or infrequently accessed ISAM data sets.

Migration

A Mass Storage System user must meet the VS prerequisites and decide on his conversion priorities. He must install an MSS and have installed at least one 145, 155II, 158, 165II, 168, 158 Multiprocessor or 168 Multiprocessor, the appropriate OS/VS1 or OS/VS2 release, and the Virtual Storage Access Method (VSAM). If he has other VS conversions such as IMS, CICS, TSO, or JES3, he must decide on priorities or determine whether he has sufficient resources for parallel installations.

Between ordering a Mass Storage System and installing it, there are several activities which will ease the conversion if they are completed prior to hardware installation. These concern non-labeled tape and non-cataloged data sets, JCL usage, reblocking, existing device dependencies, OS/VS device independence, migration to MSS, and multiple CPU migration. The following paragraphs describe them.

Non-Labeled Tape

Data sets converted to 3330 or Mass Storage volumes should be labeled. The labeling activity can be completed prior to the MSS installation.

Standard labels are recommended to ease conversions and eliminate modifications. Labeling is also recommended for tape data sets which will not be initially converted, such as low activity, backup, interchange, security, archive, etc. Labeling encourages data set naming and improved system control. In this regard, it is desirable to establish a data set naming convention, if none already exists. Aside from providing future guidance to application programmers, this practice should eliminate duplicate data set names, which are not acceptable when cataloging data sets.

Non-Cataloged Data Sets

Case studies have revealed that, in some instances, a tape data set may be named but not cataloged. Cataloging is recommended, but not required, except for data sets which are to be converted to VSAM.

The VSAM catalog is an option for OS/VS1 and OS/VS2 Release 1 and a requirement for OS/VS2 Release 2 and subsequent releases. Its availability makes it possible for the user to plan and complete this recommended activity prior to the Mass Storage System installation. Cataloged procedures, with installation standards, improve the management of data because the JCL is in one location, standardized, and under system control.

JCL Usage

An installation without JCL standards probably has a variety of practices in use which make interpretation and conversion difficult by either manual or automated methods. Analysis of JCL usage, and adoption of standard practices is a pre-installation activity that simplifies conversion to the Mass Storage System.

In an IBM internal installation, analysis of JCL revealed 10-15 variations in writing JCL for some functions. Of some 22,000 DD statements examined, about eight percent could not be automatically classified as tape or DASD. Because JCL varied with each programmer, interpretation of JCL without actual execution was difficult.

Complete device independence is advantageous. Use of group or symbolic UNIT names, the addition of the SPACE parameter (ignored for tape data sets), implementation of standard IBM JCL practices, plus the previously described pre-installation activities, minimize the effort required to make the MSS conversion.

Reblocking

Consideration should be given to blocking data sets at 4,000 bytes per block or greater. Performance degradation will result from blocking data sets at less than 4,000 bytes.

Existing Device Dependencies

You can have programs with device dependent code, fixed block sizes, or error exit routines that complicate your conversion to the Mass Storage System. The effort to redesign these programs can be minor or major, depending on the application, the program documentation, and the availability of programmers who are familiar with them.

A conversion decision must be made on these programs. Those using a large number of tape drives should be converted. Those that can operate with the tape drives to be retained need not be converted initially.

Many users have been striving to eliminate device-dependent code for years. With a sound pre-installation plan, the programs to be recoded can be identified, converted to DASD, and tested prior to the installation of the Mass Storage System.

OS/VS offers device independence to the user via the BSAM and QSAM access methods. A user planning to install MSS can alternately access tape, DASD, or the Mass Storage System (DASD) data sets by using the appropriate BSAM or QSAM macros.

BSAM		QSAM
(1)	(2)	(1)
READ	CONTL	GET
WRITE	BSM	PUT
BSP	FSM	PUTX
OPEN	BSR	OPEN
POINT	FSR	CLOSE
CLOSE	RDBKW	
(1) Tape, 3330, or MSS (3330V) (2) Tape only		

The use of group or symbolic UNIT names and the addition of the SPACE parameter simplify the media transition.

Migration to the Mass Storage System

A representative user can migrate to the MSS starting with his tape library. This area is among his most expensive operations and improvement promises immediate savings and a relatively short payback on the investment in the Mass Storage System.

Before MSS is delivered, he completes installation of his VS hardware, the appropriate OS/VS operating system, and VSAM. His tape library should be labeled and cataloged, and his tape programs should be device independent (at least this effort should have been completed for tape data sets destined for the MSS).

He has also reviewed and changed his retention, backup, off-site security, and archive procedures. To implement these changes, he uses generation data groups and more system control.

Just prior to installation, he is primarily concerned with identifying the jobs and data sets that tie up the greatest percentage of his system and operation resources, scheduling their conversion in a priority sequence, and managing the conversion so that all jobs using each data set are identified and converted.

After installation and checkout of the Mass Storage System is completed, the user initializes the Mass Storage Control tables with his configuration and data paths and executes selected jobs to thoroughly test the system. He runs a full in/out cycle and checks the results to make certain that the system is functioning correctly for input and output.

When the initial testing is completed, the conversion begins on a production basis. Each day, the JCL for selected data sets is changed and the jobs are run on a "tape or DASD in/MSS out" basis. Using a "downstream approach", the "MSS" data set is cataloged so that the next execution will be "MSS" in "MSS" out. The catalog contains the data set name and identifies the unit as 3330V.

Key jobs are scheduled earlier in their normal cycle to allow sufficient time to check the output and rerun if required. Minor jobs are not tested unless unusual system difficulties or errors are encountered.

There are a few data sets that cannot be converted conveniently as "tape in/MSS" out. For these, a tape to Mass Storage System (tape to DASD) copy or a DASD to Mass Storage System convert is required. These include tape masters updated monthly or less frequently, tables, backup, and special tapes. The copy is by data set except for the 3330 Model 1 or 2, which can be copied by volume.

Multiple CPU Migration

The user with multiple CPUs parallels the single CPU migration up to a point. When his Mass Storage System is installed and checked out, he initializes the Mass Storage Control tables with his multiple CPU configuration but begins testing with only the primary CPU and a single Mass Storage Facility. The unused devices are varied off until they are connected and ready for test.

As in the single CPU installation, selected jobs should be run and the results checked to make certain that all components are functioning correctly. When this phase is completed, the second CPU can be attached and selected jobs run again, this time testing the attributes of SHARE and EXCLUSIVE.

Components are added to the system one at a time until the complete system is operational. Conversion then proceeds on a production basis until complete.

The start-up time is longer for the multiple CPU user because there are more components and more paths to test. His production conversion time should be proportional to the number of data sets to be converted.

Conversion Plan

There are many interrelated items for consideration and review in determining the best method of conversion. Initially, one can think in terms of applications or important jobs which use major system resources as the starting point. The data sets and the combination of job/data set interdependencies determine the final conversion sequence.

Refer to the *IBM 3850 Mass Storage System Installation Guide,* which contains a more detailed explanation of the installation and conversion process. The interrelationships which affect conversion become more complex as your data base becomes more integrated. As this occurs, it becomes increasingly important to schedule the conversion by data set.

Physical Installation and Test

This section describes the installation process by the IBM installation team. For presentation here, it is divided into three phases:

Phase one includes the placement, assembly, conversion (3830 Model 2 to Model 3), and cabling of the Mass Storage System components. A carefully checked physical planning layout is essential. See the *IBM System/370 Installation Manual - Physical Planning* (GC22-7004 for U.S. users or GC19-0004 for World Trade users). As the numbers of CPUs, 3830 Model 3s and MSFs increase, the distance between the machines becomes more critical because of multiple paths and cable length restrictions. The following table shows **maximum** cable lengths (x dimension).

Location	Feet	Meters
Mass Storage Facility to Byte or Block Multiplex Channel	200	61
3830-3 or ISC to Block Multiplex Channel	150	46
The furthest 3803-3 or ISC to a Mass Storage Control (MSC Table Packs)	150	46
The furthest 3803-3 or ISC to a Mass Storage Control (non-MSC Table Packs)	300	91
3333 to 3830-3	200	61
Data Recording Control to 3830-3 or ISC	200	61
3333 to ISC	200	61

Testing at this phase is by component using the built-in microdiagnostic programs controlled from the CE panel. During this phase, 3830 Model 2 Storage Controls can be field converted to Model 3 and the Staging Adapter can be field installed on the 158-168 Integrated Storage Control feature(s). These field conversions can be made in advance of the physical installation of the Mass Storage Facility.

Phase two consists of a 3830/3333/3330 subsystem test, installation and checkout of the Mass Storage Facility, and preparation to activate the Mass Storage System by initializing the MSC tables and staging packs. The 3830 Model 3 can operate in virtual mode (staging) or in normal mode (non-staging). Phase two testing and initialization is done in normal (non-staging) mode.

The initialization of the MSC tables must be done on a non-staging drive. A special testing configuration oriented to the diagnostic routines, rather than the user configuration, is loaded at this time. When the MSC Table Packs are moved to their primary location and the MSC initialized, these tables are accessible only by the MSS.

All staging packs, whether previously initialized for 3830/3330 operation or new, must be initialized using a modified utility (IEHDASDR). This is necessary because the MSC requires a special format for label and VTOC and uses a modified alternate track assignment.

Phase three is a system test of the Mass Storage System. The MSC table packs must be moved to their assigned locations (or address plugs changed) and the 3830-3 changed so that all specified 3830 Model 3s operate in a virtual mode. An Initial Microprogram Load (IML) causes the MSC to read the configuration tables on the MSC table pack, verify the existence of each device, and check the addressing. If all tests are successful, the MSC loops, awaiting further instructions.

The installing Customer Engineer can now insert special cartridges and proceed with his system tests. As the cartridges are inserted via the cartridge access station, the MSC causes them to be moved to data recording devices, reads their labels, recognizes them as CE cartridges, places them in the cartridge store, and updates its inventory to reflect their presence. At the end of this phase, the Mass Storage System is ready for customer use.

Performance

Performance of the 3850 Mass Storage System depends on a complex set of hardware and software variables as well as system programming and system design considerations. All of these factors influence some aspect of performance in an environment which includes a Mass Storage System.

At the central processor level:

1. Amount of CPU storage

2. Number of job steps running concurrently

3. Type of system control program and the mode of its use

4. Number and type of CPUs

5. Number of available channels

6. Skill of the console operator

At the Mass Storage System level:

1. The number of data paths between the Mass Storage Facility and the Storage Controls as determined by the model number of the MSF.

2. The number of 3830-3 Storage Controls and 3333/3330 Disk Storage Drives available to the Mass Storage System for "buffer storage".

3. Data set size.

4. Block size.

5. Allocable space within a staging drive group.

6. Use of volume attributes (for example, DASDERASE).

7. Use of staging parameters (for example, BIND).

8. Access methods and catalogs.

9. DASD file organization (for example, small data sets should not overlap cylinder boundaries).

10. Data reuse

11. Data set location within the mass storage volume.

In a batch environment, the following factors favorably influence total system performance when the 3850 MSS is installed:

- Manual handling of tape reels and disk packs is reduced.

- Operating system allocation problems are significantly reduced.

- Increase in availability of data reduces job scheduling problems.

In an interactive or DB/DC environment, system performance is minimally affected by the inclusion of the 3850 MSS, if the same configuration is dedicated to DB/DC. This requires:

- The same number of BOUND volumes as previous real volumes.

- The same path arrangement (that is, same number of BOUND volumes per 3830 Model 3 as previous real volumes per control unit).

- The same switching arrangement (string switches).

- The data set must be pre-staged prior to terminal access and destaged after terminal activity has ceased.

- No other staging activity can take place on these paths.

- The same data organization must exist on BOUND volumes as previously existed.

Your IBM Marketing Representative and Systems Engineers are educated in the use of a series of new IBM Aids which require as a principal input records 14 and 15 of Systems Management Facilities (SMF) data. These new tools assist in quantifying the amount of data processed currently for individual or groups of jobs and help you to configure the Mass Storage System. Your IBM account team, working with special support groups, will assist you in defining and configuring a Mass Storage System for your computing environment.

Financial Justification

The Mass Storage System offers many users two important attributes:

1. It is an IBM offering that opens new vistas of application development because of the large amount of relatively inexpensive storage under system control.

2. It can provide an opportunity for many users to realize operational savings soon after installation. These savings could be invested in applications development.

In support of the development of the Mass Storage System, several case studies were conducted to identify the kinds of applications that the Mass Storage System made economically feasible for the first time. Other case studies were done to identify in detail the areas of operational savings and amounts.

New applications are those that require very large online storage with moderate response requirements. For example:

- In a major bank, a microfilm inquiry bureau has been established to answer requests against current customer accounts. Each month, as statements are printed, a microfilm copy of each statement is retained. When a customer requests information about his account through a branch bank, the inquiry is forwarded via terminal to a computer center where an index is found on DASD and forwarded with the original request to the microfilm bureau. The bureau then obtains the requested record(s) and answers the question either by teleprocessing (urgent) or by mail. In the study team's opinion, if the entire application were placed on a system utilizing IBM's Mass Storage System, customer service would be improved and the savings to the bank would be substantial.

- In a large manufacturing company, a data collection application has been implemented manually to obtain and correlate information concerning its industry. Presently it takes one day to one month to respond to a request from any of several hundred users of the data, and sometimes a second or third request is necessary. Using the IBM program product *STAIRS* to maintain the data on the Mass Storage System, the study group concluded, would save up to ten percent of the professional user's time, amounting to very large savings that would improve the effectiveness and decrease the cost of this function.

- One large insurance company is preparing and changing all of its automobile policies manually and updating computer files by batch runs at night.Competitors are offering better client service through online policy creation, updates, and claims service. The company could not move to automating these very large files, however, because of the prohibitive cost of sufficient DASD. With MSS, the entire file could be placed under system control resulting in large savings and better customer service.

- A government agency provides a service to the populace that involves inquiries to a tape file of approximately 240 billion bytes. The inquiries are complex, the file diverse, and the time taken to process requests with the present semi-automated approach can be in excess of a year. In the opinion of the people who conducted this study, if the data were captured under system control in the Mass Storage System, the response time could be reduced to

one or two weeks. No value estimate was made for this application except that it represents an improvement in a critical and visible service to millions of voters.

- In a U.S. bank, the adjustment history file represents one year's processing of all MICR (checks, deposits, etc.) transactions. This file is contained in approximately 27 billion bytes and would be the equivalent of 200 disk packs of data. The post cycle adjustment process at this bank requires more than 60 clerks and a considerable volume of paper and card files. The adjustment department handles approximately 300 adjustments per day, or about five per clerk. Adjustments can be found for documents which were distributed but not accounted for, accounted for but not distributed, misread by the MICR machines, MICR encoded incorrectly and posted to the wrong accounts. With the Mass Storage System to store and manage the large adjustment history file and IMS to build a transaction-oriented application, the team estimated that the manpower requirements to process these adjustments could be sharply reduced, yielding a savings of about twenty thousand dollars per month. Additional large dollar savings were identified with the reduction of float and adjustment write-offs. Performance of the Mass Storage System was judged to be adequate to do those applications that the bank would convert to it, plus this one very important application.

There are mass storage applications like these in most industries. The study teams concluded that the users studied would install the Mass Storage System because they had problems that they wanted to solve that cannot be solved without a mass store. In almost every case, too, there existed solid financial justification after the estimated cost of development of the new or expanded application.

Three areas of potential **operational savings** have been identified: tape libraries, computer rooms, and computing systems. While there is no attempt here to estimate through "rules of thumb" savings in these areas, they are provided for reference. In tape libraries studied, there appeared to be potential savings from:

1. A reduction in the number of reels of tape and disk packs that were routinely scheduled to be replaced.

2. A reduction in the time and resource allocated to tape cleaning and certification.

3. A reduction in library record keeping and occasional errors such as use of the wrong physical label.

4. Better data security than with conventional tape and DASD.

5. A reduction in required floor space.

6. A potential reduction in the labor content for volume handling, with an attendant reduction in training and management time.

In the computer room, the possible savings noted came from these areas:

1. A reduction in the resources required to update run sheets.

2. A reduction in the amount of JCL change activity after fully automating volume handling.

3. A reduction in console operators' time taken to communicate with tape handlers.

4. A labor reduction with an attendant reduction in management time and personnel training.

5. A requirement for fewer tape drives.

Within the system, potential increased efficiency was observed from these sources:

1. Improved system productivity through the sharp reduction in contention for I/O devices available.

2. Faster job turnaround from the elimination of wait time for manual volume handling.

3. Elimination of mount delays, mount errors and some operator errors.

4. Better utilization of DASD and the remaining tape drives.

Operations savings tend to be more easily quantified than the potential economies from new applications, and they can usually be achieved faster. However, the studies show that the MSS can provide operations savings **and** new applications in the same installation.

Glossary of New Terminology

With the announcement of the IBM 3850 Mass Storage System, several new terms are being introduced to the industry. Some of these are listed here for your reference.

accessor: This component of the *Mass Storage Facility* selects cartridges from the *cartridge store* and transports them to the *data recording devices*. It also moves cartridges from the data recording device to the cartridge store, and to and from the cartridge access station. There are two in each Mass Storage Facility.

accessor control: Each MSF has duplexed accessor controls to decode and sequence messages from the *Mass Storage Control*.

acquire: The process of allocating space on a *staging drive* and *staging data* from a *data cartridge* to a *staging pack* on that drive, when required.

cartridge access station: Openings on the front of the *Mass Storage Facility* to permit manual entry and exit of cartridges.

cartridge label: An area on the tape which contains information about the cartridge.

cartridge store: That part of the *Mass Storage Facility* consisting of the *cells, accessor,* and *accessor controls.*

cells: Data cartridge storage locations within the *Mass Storage Facility.*

cylinder fault: A "cylinder fault" results from the issuance of a SEEK command for data that has not been staged.

data cartridge: Refers to a cartridge.

data recording control (DRC): Each pair of *data recording devices* has a primary data recording control. This unit encodes and decodes data recorded or read, controls tape movement and assists in error recovery. In Models 2, 3 and 4 each data recording device also has an alternate data recording control.

data recording device (DRD): Reads from and records information on the tape in data cartridges.

destage: The process of transmitting data from DASD to a *data cartridge.*

extended group coded recording (E/GCR): Refers to the technique used to encode data on a *data cartridge.* It includes error correction code.

Mass Storage Control (MSC): This component of the *Mass Storage Facility* controls functions for the entire *Mass Storage System.*

Mass Storage System Communicator (MSSC): That part of the VS1 or VS2 system control program which communicates with the *Mass Storage Control* part of the *Mass Storage Facility.*

mass storage volume (MSV): Two logically associated *data cartridges.* Collectively, 100 megabytes of storage capacity on data cartridges in 3336 Model 1 format.

mass storage volume control functions (MSVC): Name given to a collection of functions which are designed to assist the space manager in managing space on *mass storage volumes.* These functions reside in the *Mass Storage System Communicator.*

mass storage volume inventory data set or inventory data set (MSVI): Refers to the data set containing information about the *mass storage volumes.* This data set resides on DASD and is accessible through the host CPU(s).

primary host: The host CPU in a multi-host system configuration that has the responsibility of processing unsolicited interrupts from the *Mass Storage Control.*

Space Manager: The person who is responsible for the management of space on *mass storage volumes.*

staging: The process of transmitting data from the *data cartridge* to a *staging drive.*

Staging Adapter (SA): Refers to the feature on a System/370 Model 158 or 168 Integrated Storage Control (ISC) which equips it to operate in a *Mass Storage System.*

staging drive: A DASD drive allocated at systems generation to receive data from a *Mass Storage Facility.* Can be a 3333/30 Model 1 or 11 or a 3330 Model 2.

staging drive group: A logical association of *staging drives* created by the MSC Table Create Utility for the purposes of space management and recovery.

staging pack: A 3336 Disk Pack which has been initialized to receive data from a *Mass Storage Facility.* Can be a Model 1 or 11.

virtual drive: A unit control block to which one *virtual volume* is assigned by a host operating system and which relates to a *staging drive group.*

virtual volume: Refers to staged data from a *mass storage volume.*

3830 Model 3 Storage Control: Refers to the model of the 3830 DASD control equipped to operate in a *Mass Storage System.* Supports staging and non-staging drives.

3850 Mass Storage System (MSS): Describes the entire storage system including the *Mass Storage Facility* and the connected DASD subsystem (3830 Model 3 Storage Controls, 3333 Disk Storage and Control, 3330 Disk Drives, and the Staging Adapter feature on 158/168 Integrated Storage Control).

3851 Mass Storage Facility (MSF): Refers to the cartridge storage facility and includes the *accessors, data recording devices, data recording controls,* and their associated support systems and one or two Mass Storage Controls.

Bibliography

Besides this manual, you can read these IBM publications for more information about the IBM 3850 Mass Storage System. They will be available through your IBM Branch Office by the times indicated.

TITLE	DESCRIPTION	AVAILABILITY
IBM System/370 Installation Manual - Physical Planning (GC22-7004-2 and later editions for U.S. users and GC19-0004-2 and later editions for World Trade users)	This document discusses physical planning considerations such as physical dimensions, air conditioning requirements, cable lengths, etc.	At announcement of the MSS.
OS/VS Mass Storage System (MSS) Planning Guide	This publication contains early planning information for the Computing Center Manager and Systems Programmer with emphasis on software support of the MSS.	Before first customer ship
OS/VS Mass Storage System (MSS) Utilities for MSS Space Management	This document contains information on how to use the new mass storage volume space management utilities and functions of the MSSC.	Before first customer ship.
OS/VS Mass Storage Control Table Create Utility	This document describes the use of the MSC Table Create Utility.	Before first customer ship.
OS/VS Data Base/Data Communications Mass Storage System (MSS) Planning Guide	This document describes configuration, processing and performance considerations (MSS and DB/DC) when MSS is introduced into a DB/DC environment. It includes examples of possible MSS usage in an IMS/VS or CICS/VS environment.	Before first customer ship.
IBM 3850 Mass Storage System Principles of Operation	This document describes the interaction among the components of the Mass Storage System and includes some error recovery information.	At first customer ship.
OS/VS Operators Library: Mass Storage System (MSS) Reference	This publication describes operating instructions and provides some error recovery information.	At first customer ship.
IBM 3850 Mass Storage System Installation Guide	This document contains detailed installation management and conversion information based on several early installations.	At first customer ship.

In addition, existing IBM publications which are affected by the announcement of the 3850 Mass Storage System will be updated to reflect this product.
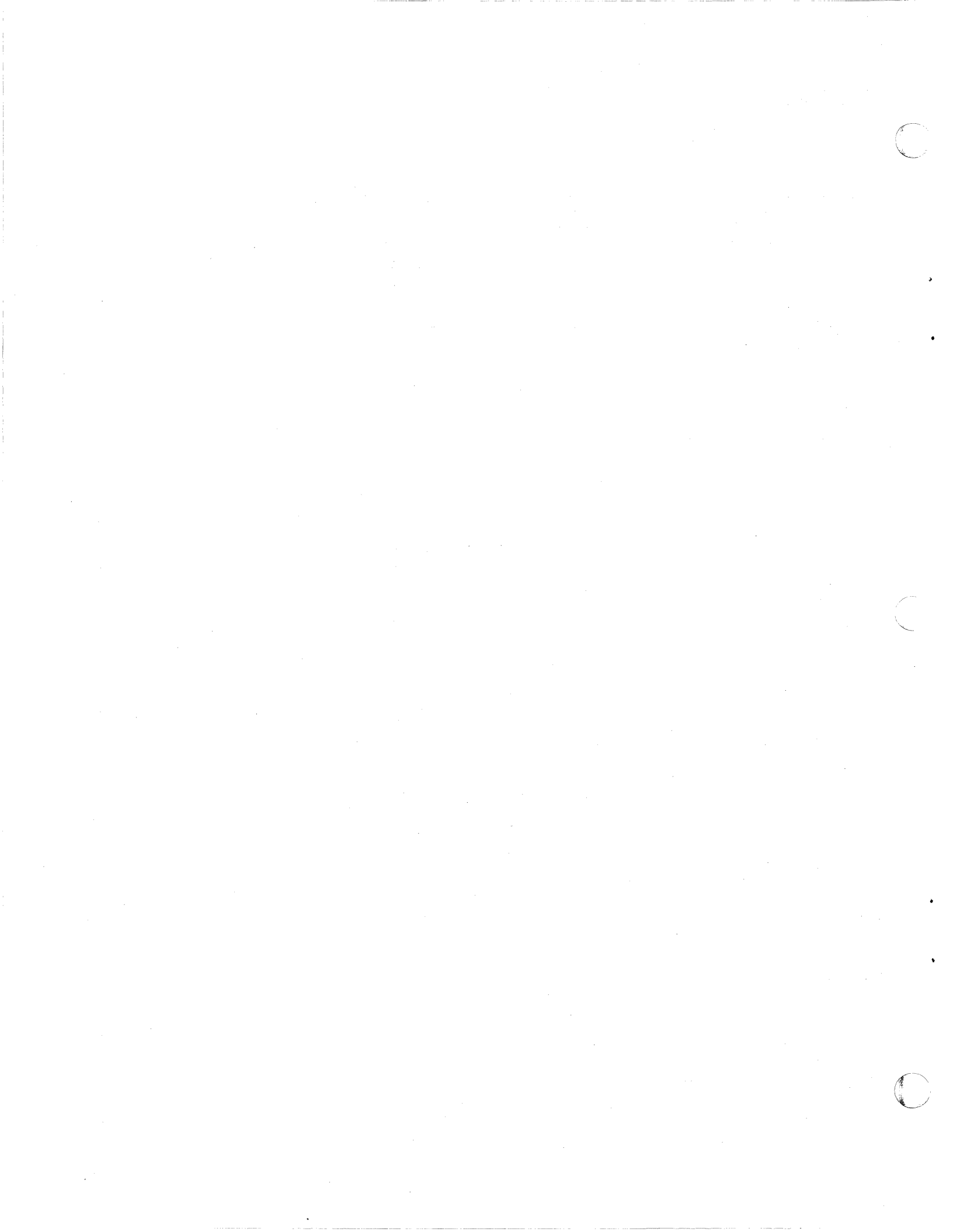
Index

The entries in this index are shown as they appear in the text of the book, that is, capitalized words are capitalized in text, lowercase words are lowercase in text, etc. When more than one reference is given for an entry, the primary reference is listed first. Italics are used to denote figure references.

Your Comments, Please . . .

Introduction to the IBM 3850
Mass Storage System (MSS)

GA32-0028-1

Your comments and constructive criticism regarding this publication will help us to improve it so that you may be better served. Each comment will be reviewed carefully by those responsible for the publication. Comments and suggestions become the property of IBM.
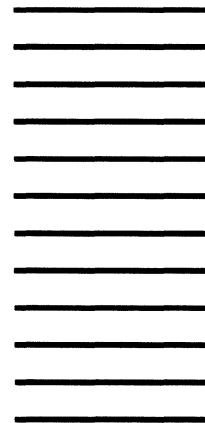
Thank you for your cooperation.

Comments:

GA32-0028-1

Business Reply Mail
No Postage Necessary If Mailed In The U. S. A.

Postage will be paid by:

IBM Corporation
Post Office Box 1900
Boulder, Colorado 80302

Attention:
MSS Support Center

Printed in U. S. A. GA32-0028-1

GA32-0028-1