GA32-0062-0
File No. 4300-07/S370-07

Systems

Introduction to IBM 3880 Storage Control Model 13

GA32-0062-0
File No. 4300-07/S370-07

Systems

Introduction to IBM 3880 Storage Control Model 13

IBM®

Federal Communications Commission (FCC) Statement

The following warning applies to the IBM 3880 Storage Control Model 13, used in the U.S.A.:

Warning: This equipment generates, uses, and can radiate radio frequency energy and if not installed and used in accordance with the instruction manual, may cause interference to radio communications. It has been tested and will comply with the limits for a Class A computing device pursuant to Subpart J of Part 15 of FCC Rules, which are designed to provide reasonable protection against such interference when operated in a commercial environment. Operation of this equipment in a residential area is likely to cause interference in which case the user at his own expense will be required to take whatever measures may be required to correct the interference.

First Edition (September 1981)

Changes are made periodically to the information herein; before using this publication in connection with the operation of IBM systems, consult the latest *IBM System/370 Bibliography*, GC20-0001, for the editions that are applicable and current.

A form for reader's comments is provided at the back of this publication. If the form has been removed, comments may be addressed to IBM Corporation, General Products Division, Department 61C, Tucson, Arizona, U.S.A. 85744. IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation whatever. You may, of course, continue to use the information you supply.

Preface

This publication introduces the IBM 3880 Storage Control Model 13. It is intended for use by data processing executives, system planners and programmers, IBM marketing representatives, and system engineers. The reader is assumed to have a thorough knowledge of computer systems and storage subsystems.

The subject matter is presented in two chapters:

- "Chapter 1. DASD Management Overview," discusses what the data processing industry will experience in the next few years, and then explains how the new 3880 Model 13 uses a cache to provide a sensible alternative to current DASD-management methods.
- "Chapter 2. IBM 3880 Storage Control Model 13," describes the characteristics, features, and performance of this new model of the 3880.

Related Publications

Programmers should be familiar with the information contained in:

- *IBM System/370 Principles of Operation*, GA22-7000
- *IBM 4300 Processors Principles of Operation*, GA22-7070

Additional information about the devices that attach to the 3880 Model 13 can be found in:

- *IBM 3880 Storage Control Description*, GA26-1661
- *IBM 3380 Direct Access Storage Description and User's Guide*, GA26-1664

Contents

Figures

Chapter 1. DASD Management Overview

The business world of the '80s is experiencing an increasing dependence on data processing services. The use of online, data-based systems is expanding, as are the requirements for improved performance for the end users.

The last few years have seen a dramatic drop in the cost of processors due to technological improvements. These improvements have enabled more users to have access to the power of a large processor. The constraint to improved end-user productivity is now the inability of the supporting I/O subsystems to supply data as fast as it is needed.

Because of the different rates of product evolution, there has been a steadily widening difference between the data-access speeds of processor storage and those of the attached I/O devices. See Figure 1-1. But a fast processor usually requires fast access to large amounts of data in order to be fully utilized. Thus, it may not always be possible to realize the full potential of a fast processor if the I/O devices become a bottleneck.

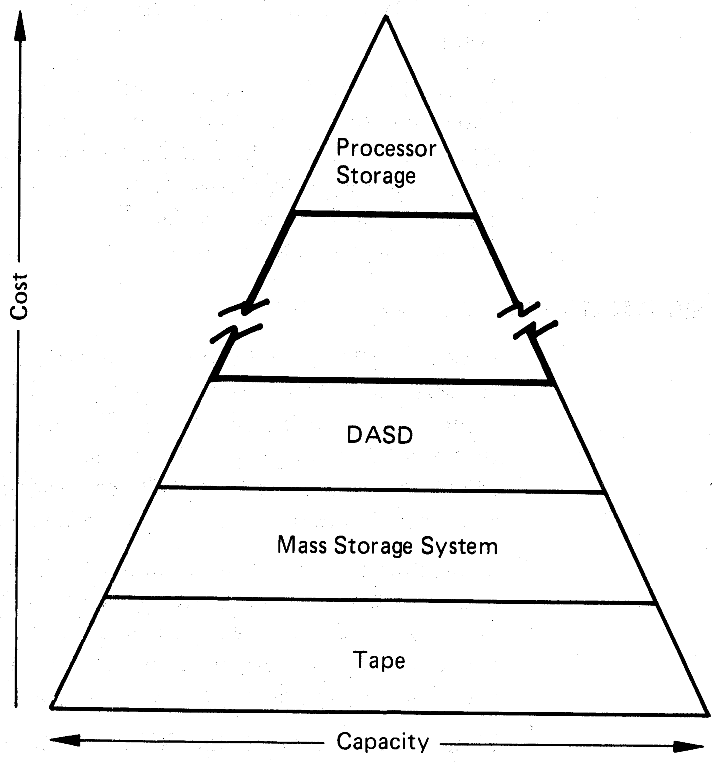


Figure 1-1. Relative Data Accessibility for Current Products

One noticeable consequence of this difference in access speeds occurs in paging. The virtual storage philosophy is that inactive pages should not occupy processor storage when they are not being referenced but should be kept in auxiliary storage until required. However, "inactive" is a relative term; the discrepancy exists between the performance of processor storage and that of the paging devices is that the processor is often constrained by the inability of the paging device to supply data when required.

In a similar manner, devices supplying application data can also introduce bottlenecks that affect system performance. To ensure that a processor is kept busy while an application waits for DASD, operating systems switch to other applications when operating in multiprogramming mode. This requires sufficient DASD actuators to match the level of multiprogramming otherwise, the processor is kept waiting. With DASD increasing in capacity and processors becoming more powerful, the number of actuators required to service the processor efficiently may well exceed the number of actuators required for capacity alone. Or conversely, the system could have sufficient DASD capacity, but the processor could be constrained because the DASD is unable to supply data as fast as needed.

To improve DASD performance requires that the mechanical delays of seeking and of disk rotation be reduced. A technique that is used in other products to improve performance is to add a buffer. For a number of years, processors have had small-capacity high-speed storage units that act as buffers between the processor and the relatively slow processor storage. Although the storage capacity of these units may only be 32K bytes, efficient algorithms manage the data so that when the processor requires data, that data tends to be in the high-speed buffer and not in the slower processor storage.

This technique of forming a storage hierarchy has been successfully implemented in other product areas using either software or microcoded algorithms to ensure that active data is kept in the faster device and inactive data is moved to the slower device. The IBM Mass Storage System (MSS) uses this principle, as does the Hierarchical Storage Manager (HSM). The former uses microcode; the latter is software.

Two Complementary Storage Controls

IBM now offers two additional models of the 3880 Storage Control that enhance the storage hierarchy. The Model 11, which is designed for paging applications, is complemented by the new Model 13, which is designed specially for nonpaging applications. Although both models reduce the differences in processor/DASD performance and improve user productivity, each model is designed to meet a specific need.

These new products bring important improvements. When businesses have a large number of users on an interactive system, variations in system performance can be measured in thousands of dollars per hour. Implementing system performance improvements that increase productivity can result in significant savings to the business.

Figure 1-2 summarizes the different applications and features of the 3880 Models 11 and 13. For further information about the 3880 Model 11, refer to *Introduction to IBM 3880 Storage Control Model 11*, GA32-0060.

3880 Storage Control		
	Model 11	**Model 13**
Designed for:	Paging and swapping	System and application data
Subsystem storage size:	8 megabytes	4 or 8 megabytes
Number of storage directors that can access the cache:	One (paging storage director)	Two
Processors to which it can attach:	3081, 3033, 3042 Model 2 3032, 3031, System/370 Models 158 and 168, 4341 Model Groups 1 and 2	3081, 3033, 3042 Model 2, 3032, 3031, 4341 Model Group 2
Channels to which it can attach:	3.0, 2.0, and 1.5 megabyte per second*	3.0 megabyte per second data streaming channels only
Type of DASD supported:	3350 for the paging storage director; 3350 and/or 3333/3330 for the other storage director	3380 for both storage directors
Operating system:	MVS/System Product Release 3	MVS/System Product Release 3
Device support:	Data Facility/Device Support	Data Facility/Device Support
*The IBM 3880 Model 11 transfers data at 1.5 or 3.0 megabytes per second only.		

Figure 1-2. Summary of IBM 3880 Models 11 and 13

Model 13 Highlights

The 3880 Model 13 is a high-performance cache DASD subsystem for use with the MVS/SP Release 3 operating system. The 3880 Model 13 has two storage directors called cache storage directors and a 4- or 8-megabyte electronic storage unit, called subsystem storage. A portion of the storage is called a cache and is used to store frequently-used application data. Both cache storage directors have access to the subsystem storage and attach the 3380 Direct Access Storage Device(s) (DASD).

The IBM 3880 Storage Control Model 13 helps to reduce the effects of the bottlenecks that often constrain DASD subsystem performance. The Model 13 provides faster access to nonpaging system and application data, and is designed to increase system throughput and reduce terminal response time, thus boosting the end users' productivity.

The 3880 Storage Control Model 13 offers:

- Fast access to disk records in the cache
- Dynamically managed cache
- Direct DASD write for data integrity
- Reduced effect of unbalanced DASD activity (DASD skew)
- Easy migration because of no user Job Control Language (JCL) changes
- No need for additional floor space than other 3880 models.
- Increased performance across attached DASD strings

Current Environment

Efficient use of computer resources includes managing DASD activity. One of the benefits of the 3880 Model 13 is its ability to manage the performance activity of system and application data. The following paragraphs examine some of the currently adopted ways of solving DASD performance problems. The term "application data" is used to mean all nonpaging system and application data that usually resides on DASD.

Application Data

DASD subsystem response time is a key indicator of DASD subsystem performance. This response time is measured from the initiation of the I/O operation by the processor until "device end" is received by the processor. This response time is a function of the time required to seek to the correct cylinder, the rotational delay in getting the requested record beneath the read/write head, the data-transfer time, and the queue time that can result from path and device contention.

A long DASD response time reduces the number of I/O instructions that a string can sustain. And the lower the number of I/O instructions, the more DASD strings and channels that are required to efficiently maintain acceptable system performance (throughput).

Several alternatives are used in an attempt to reduce the effect of long DASD response time. Here are four that are related to application data:

- Add more main storage

  This idea offers attractive options, such as increasing the multiprogramming level (MPL) and enlarging block sizes to enable transferring more data per I/O instruction. This option might be effective in sequential applications, but user JCL changes are required to take advantage of the increase in available storage. Some systems are at the maximum storage capacity; therefore, for such systems this option is not an alternative.

- Add more DASD paths

  Adding more channels, controllers, and DASD can help reduce the number of contention points, hence the queuing for DASD paths and actuators, but it can never increase performance beyond the mechanical limits of the disks themselves. Even if the processor can accommodate more DASD, planned data placement is required in order to take advantage of the extra actuators. However, this can be an expensive alternative.

- Install fixed heads

  Although data under fixed heads can be accessed without any seek time, the rotational delay and data-transfer rate are the same as for data under movable heads. The attributes of fixed heads are desirable, but many high-access data files are too large to fit on the available space.

- Tune the DASD subsystem

  Tuning a DASD subsystem requires measuring where I/O activity is occurring and moving data sets in an attempt to spread the activity as uniformly as possible across the contention points (channels, controllers, strings, and individual actuators). Effective tuning is not only difficult and expensive, in terms of professional resources, but because access patterns change as daily work changes and as new applications are installed, it is a never-ending process.

Summary

It should be clear, from the preceding discussions, that to meet the demands of the '80s, other product alternatives are needed for managing application data activity. The alternative should be a product that exhibits low service time, but doesn't have the associated capacity limitations of fixed-head DASD nor the cost associated with adding more storage devices, channels, and processor storage.

A high-speed cache, located within a DASD storage hierarchy, achieves this objective. Active data is automatically retained in the cache; inactive data in the cache is deleted. And all this activity is accomplished with a minimum of additional processor or channel overhead.

The new IBM 3880 Storage Control Model 13 uses this principle of a cache in a storage hierarchy to combine the best features of two existing technologies—the high speed of processor storage and the low cost of DASD storage.

Total Storage Management

As the demand for data processing services increases, so does the need for more data storage. The requirements for increased performance, greater system availability, and more cost-effective storage capacity, all within physical space constraints, must be met. Also, the expanding number of data sets and devices has compounded data processing management decisions in the areas of performance and space management. Meeting these needs requires a Total Storage Management approach. This approach includes assistance in selecting the best mixture of storage devices and programs to plan and manage the data in order to obtain the most effective use of storage.

The IBM 3880 Model 13 addresses the area of DASD performance and is designed to integrate with the other products that are part of IBM's Total Storage Management program.

As shown in Figure 1-3, the IBM 3880 Model 11, and now the Model 13, decrease yesterday's performance difference, and, with the other components of Total Storage Management, provide additional solutions for today's storage management tasks.

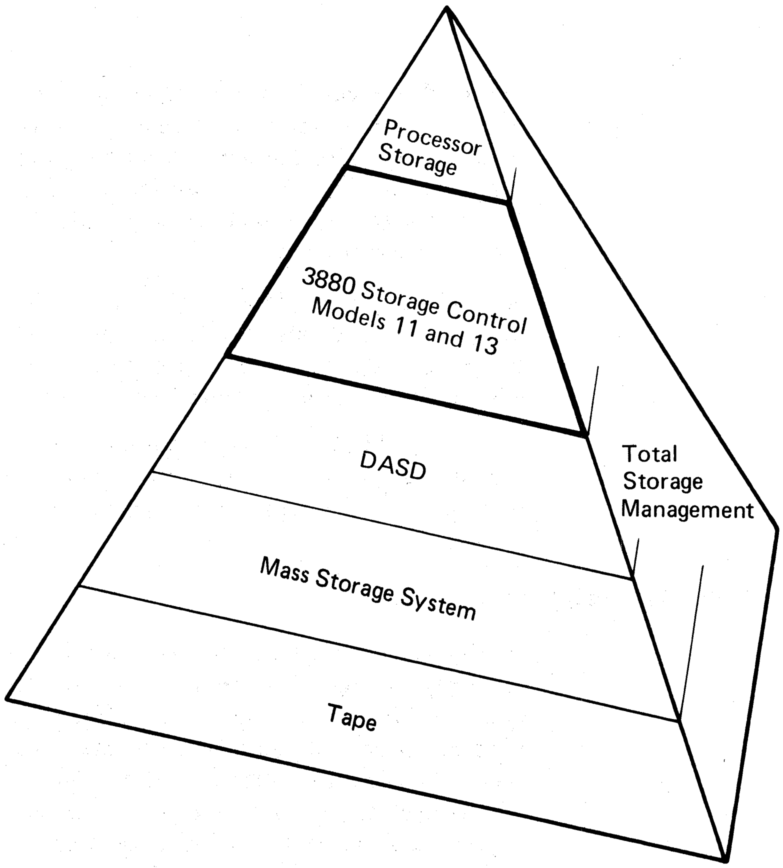


Figure 1-3. Total Storage Management

Chapter 2. IBM 3880 Storage Control Model 13

The IBM 3880 Storage Control Model 13 is a high-performance cache DASD subsystem for use with the MVS/SP Release 3 operating system.

The 3880 Model 13 has two storage directors called cache storage directors, and a 4- or 8-megabyte high-speed, electronic storage unit called subsystem storage. A major portion of this storage, called a cache, retains active DASD data for quick access. The remaining portion is called the directory and is used to locate data that is stored in cache. Both cache storage directors have access to subsystem storage. This additional storage in the Model 13 creates a two-level storage hierarchy when used with the IBM 3380 Direct Access Storage Devices (DASD).

The new storage control employs an algorithm to manage cache contents. The objective is to retain active copies of DASD records in the cache because the processor can retrieve data faster from cache than it can from DASD.

Major Components

Figure 2-1 shows the major components of the 3880 Model 13: the two cache storage directors, and the subsystem storage unit. Both cache storage directors have access to subsystem storage.

Cache Storage Directors

Each cache storage director attaches to 3.0 megabyte per second data streaming channels and strings of IBM 3380 Direct Access Storage. The Dynamic Path Selection models of the 3380 are required.

The cache storage director manages the flow of data by:

- Copying data records from DASD to the cache and from the cache to DASD
- Copying data records from the cache to the host system
- Determining which data to delete from the cache
- Searching the directory for requested data records when it is addressed by the processor
- Performing error recovery procedures (ERPs) within the subsystem storage

Subsystem Storage

Subsystem storage is offered in two capacities, 4 megabytes and 8 megabytes, and is divided into two parts. The larger area, called the cache, is divided into track slots that store data from the 3380 tracks. The smaller area is a directory that contains entries that allow data to be located in the cache.

The cache is dynamically managed by an algorithm; high activity data is accessed from the high-performance cache and low activity data is accessed from less-expensive DASD storage.

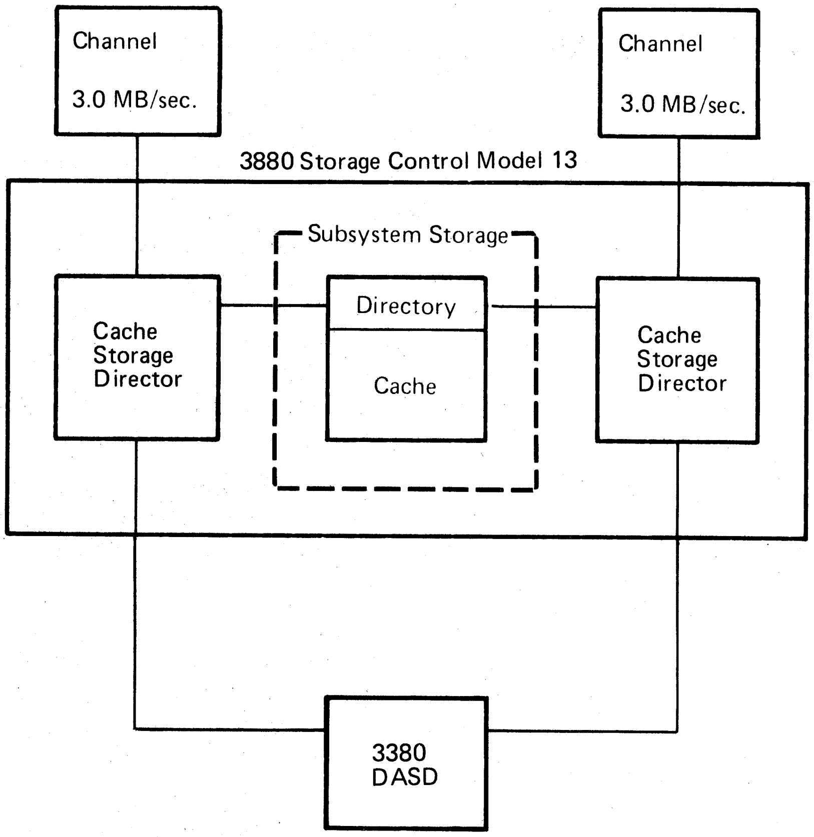


Figure 2-1. Major Components

Features

The following features are available for the 3880 Model 13:

- Two Channel Switch—Pair
- Two Channel Switch—Pair, Additional
- Remote Switch Attachment
- Remote Switch Attachment, Additional

Two Channel Switch—Pair

This feature permits each cache storage director to connect to two different channels. A single channel cannot attach to both cache storage directors. See Figure 2-2.

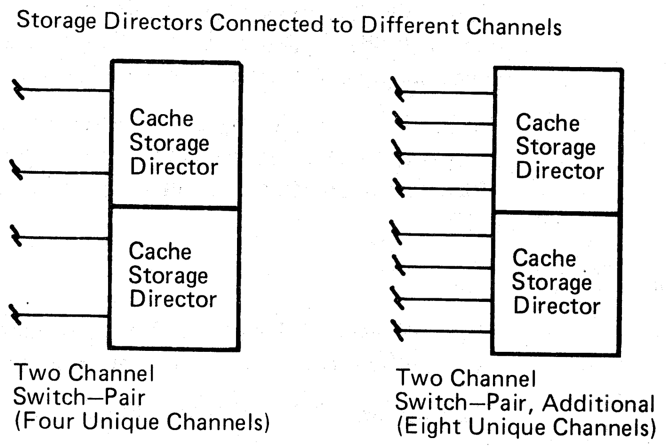


Figure 2-2. Channel Switch Options

Two Channel Switch—Pair, Additional

This feature permits each cache storage director to connect up to four different channels. A single channel cannot attach to both cache storage directors. See Figure 2-2.

Remote Switch Attachment

This feature relocates the channel enable/disable switches from the 3880 Model 13 operator panel to a remote location, usually on the processor's operator panel. This enables the operator to reconfigure at a central point rather than going from unit to unit to enable or disable the desired storage director.

Remote Switch Attachment, Additional

This feature relocates the four additional channel enable/disable switches provided by the Two Channel Switch—Pair, Additional feature from the 3880 operator panel to a remote location.

Programming Support

The 3880 Storage Control Model 13 is supported by the MVS/SP Release 3 operating system, and the program product Data Facility/Device Support.

Except for programs tightly bound to device characteristics, user programs that reference data on 3380 DASD should execute without modifications or recompilation. An MVS/SP I/O system generation is not required.

The Model 13 generates information relating to the operational characteristics of the DASD subsystem. This data, when coupled with Resource Measurement Facility (RMF) reports, should give a picture of DASD subsystem performance.

The cache can be set online or offline to device addresses. This allows an actuator with a low hit ratio, as indicated by the DASD subsystem information, to be excluded from using the cache.

The Model 13 supports the requirements for Set High Performance Storage Limits (bind). When certain tracks of data, such as indexes, are needed on a continuous basis, they can be bound in the cache.

Sequential data sets may be processed by the cache in a manner different from normal cache management. When processing certain sequential data sets cache management is able to stage ahead in order to act as a look-ahead buffer.

Processor Attachment

The 3880 Model 13 can attach to the following IBM processors through the 3.0 megabyte per second block multiplexer channel:

- Processors 3081, 3033, and 3042 Attached Processor Model 2
- Processors 3032, 3031, and 4341 Model Group 2

Channel Attachment

The 3880 Model 13 requires a 3.0 megabyte per second block multiplexer channel to attach to any of the processors. The IBM 3031, 3032, 3033, and the 3042 Model 2 processors require the data streaming feature in order to support the 3 megabyte per second data transfer required for 3380 DASD. No special features are required for the IBM 3081 processor.

The IBM 4341 Model Group 2 processor can attach the 3880 Model 13 without modification if the processor was manufactured after March 1981. If the processor was manufactured before March 1981, it requires an engineering change to have the data streaming feature added.

DASD Attachment

The Model 13 attaches only 3380s in strings consisting of an AA4 model and up to three B4 models. The minimum 3380 configuration is a single AA4 model. The maximum 3380 configuration is two AA4 models with three B4 models attached to each AA4. The 3380 Model AA4 has dynamic path selection, which is required for attaching the 3380s to the 3880 Storage Control Model 13. If there are two strings of 3380s, both strings must attach to the same 3880 Model 13. The 3880 Model 13 can have up to 32 device addresses.

Figures 2-3 and 2-4 illustrate the possible configurations.

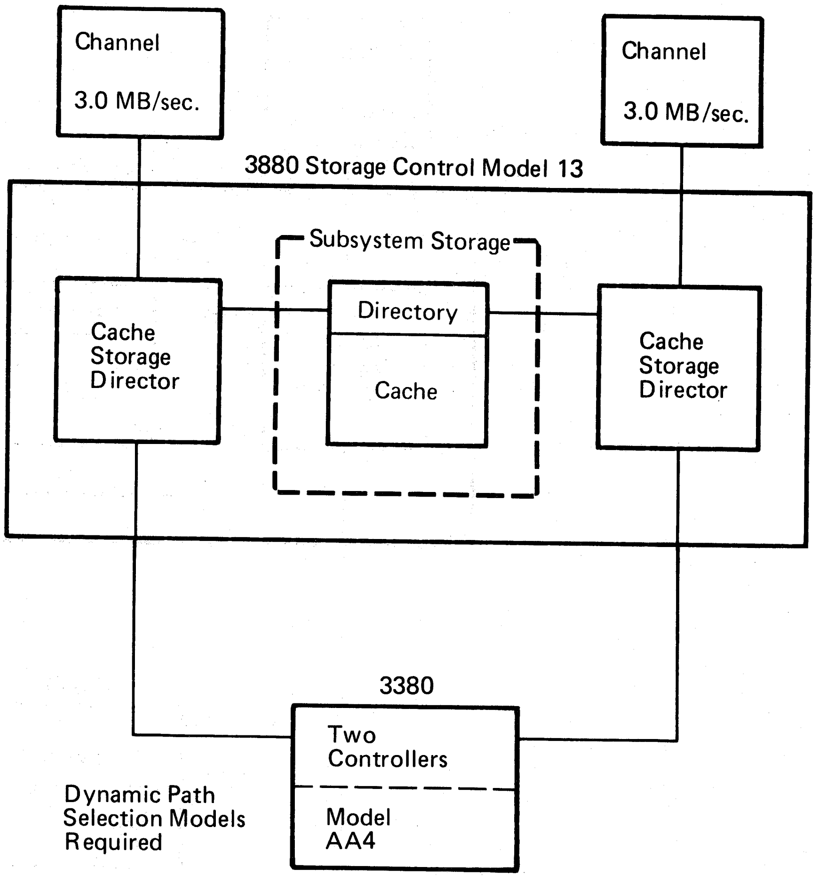


Figure 2-3. Minimum DASD Configuration

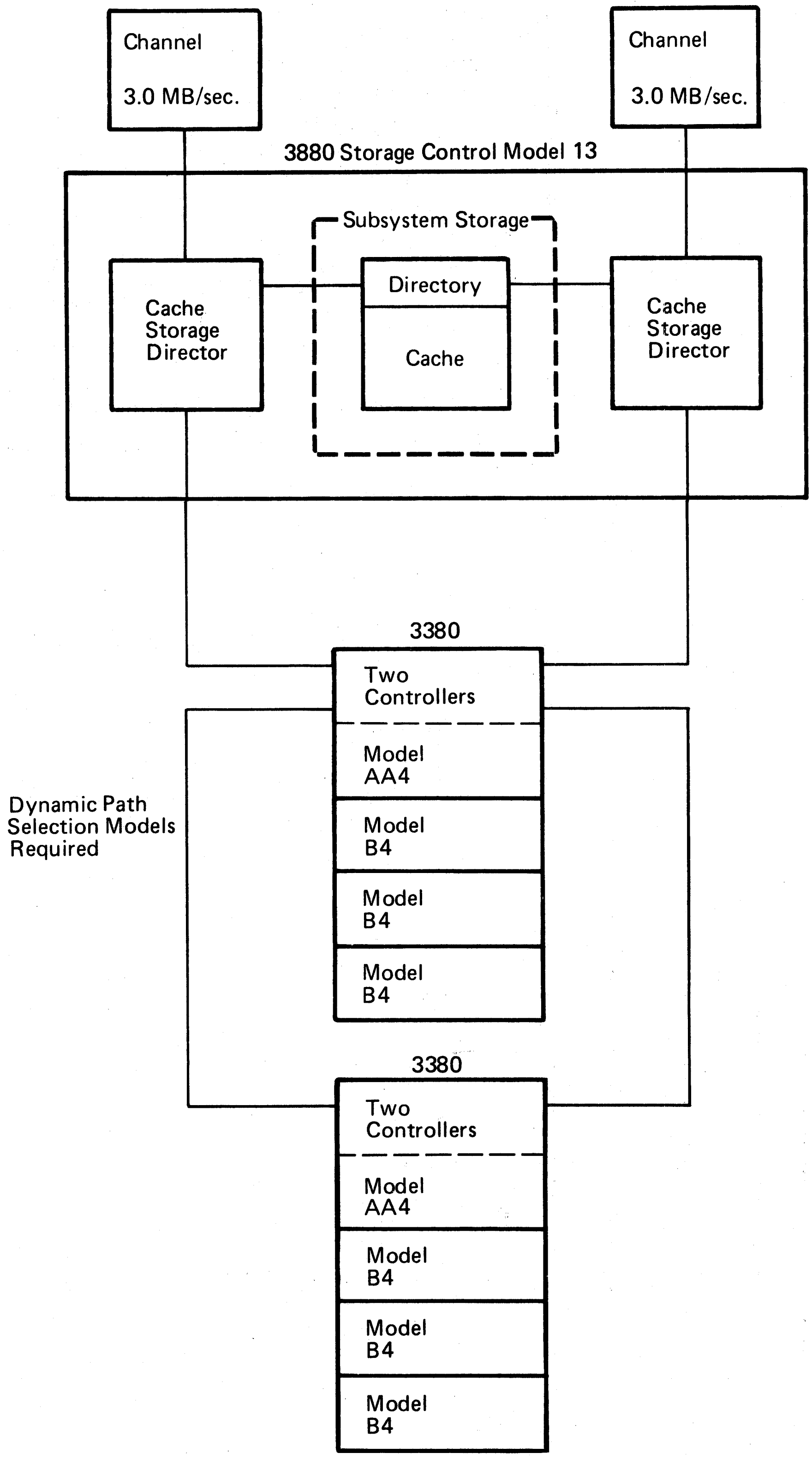


Figure 2-4. Maximum DASD Configuration

Performance

The modeled performance values in the following paragraphs show the approximate relative performance of a cache DASD subsystem and a noncache DASD subsystem. The model results that are presented used a *fixed set of assumptions in an attempt to ensure a fair and accurate comparison. However, the assumptions may not match any specific user or environment. Therefore, it is important to remember that some of the specified assumptions may not apply for a given production environment. The values shown by the performance curves should not be considered actual obtainable results in a specific environment, but rather, as an indication of the relative performance of the product.*

It is important, when evaluating DASD subsystem performance curves, to realize that overall system performance may not result from improvements in DASD subsystem throughput and response time. Other system resources, such as processor cycles, main storage size, operating system, and the communications network, may limit the potential system performance even though the DASD subsystem performance has been improved.

Performance Variables

The performance variables often associated with DASD subsystem performance are:

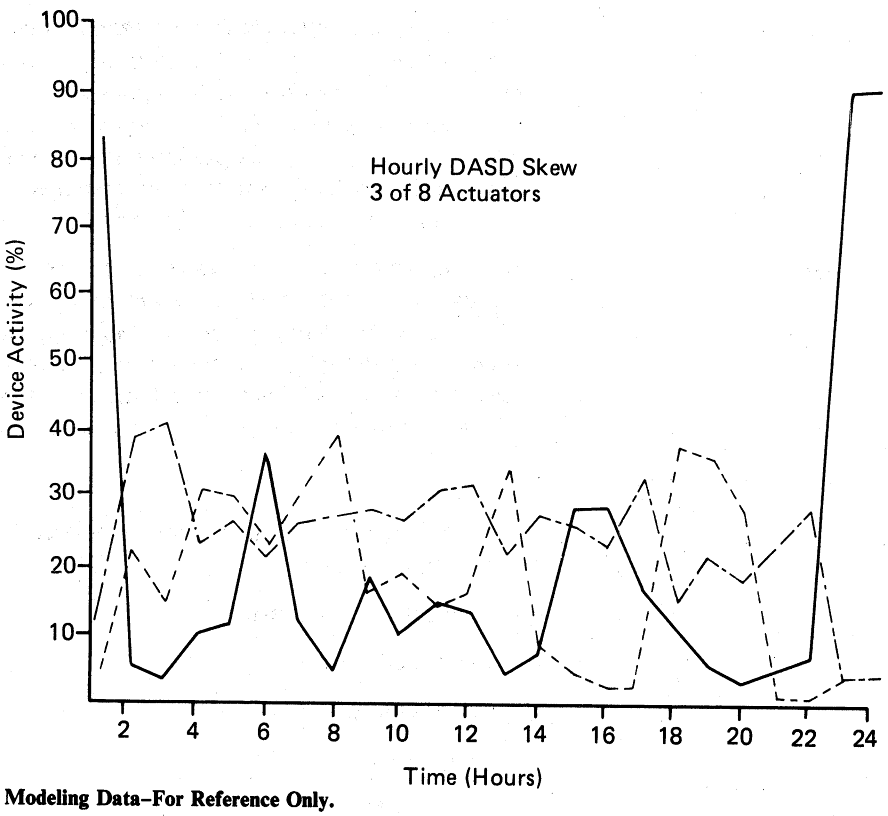
- Control unit processing
- Device seek time
- Rotational delay (latency)
- Search time
- Data transfer rate

These variables are generally adequate to compare the performance of different DASD types. But when comparing cache with noncache DASD, another variable, called *DASD skew*, is important. DASD skew can reduce the effective performance of a string of DASD. But, with a cache, the effect of DASD skew on DASD subsystem performance can be reduced.

DASD Skew

During application processing, disk accesses are *not* evenly distributed over all the devices of a DASD string. This uneven distribution of accesses is referred to as DASD skew. Skewed activity reduces the effective accesses per second possible on a DASD string. The effect of a uniform distribution tends to exaggerate the expected performance of a DASD string because no one actuator can be a bottleneck.

The usual method of overcoming DASD skew is to tune the DASD subsystem. This is difficult, and nearly impossible, because the skew pattern itself is dynamic in character; that is, the devices receiving the most activity change from time to time. Figure 2-5 shows an example of these dynamics. While the figure shows only three actuators of an eight actuator DASD string for diagram clarity, it is apparent that large fluctuations in activity take place over a period of time. Modifications to placement of data sets based on data from only one time period may have detrimental effects on performance during some other time period. This difficulty is faced each time the system programmer attempts to balance DASD activity.



Modeling Data–For Reference Only.

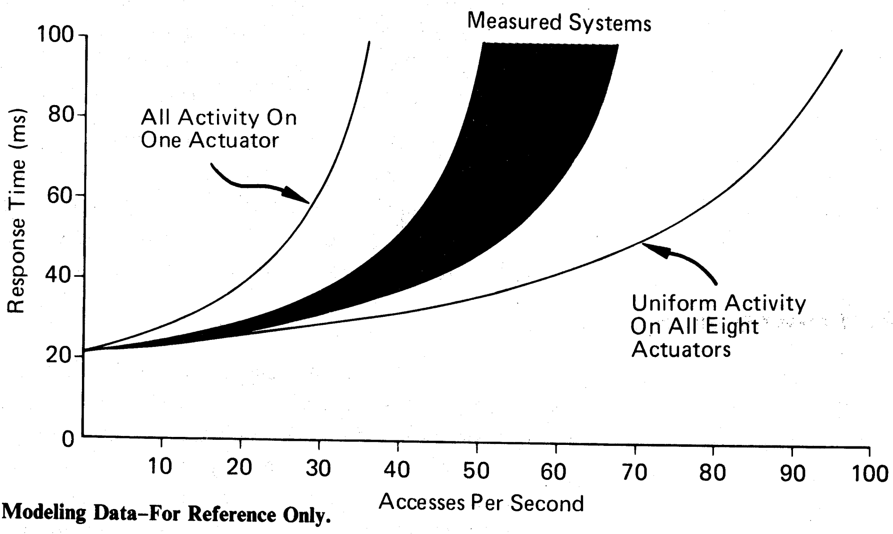
These results are from an analytic model using specific workloads and specific predefined parameters. Individual systems may fall outside these ranges.

Figure 2-5. DASD Skew

In contrast, the 3880 Model 13 helps dynamically "tune" the DASD subsystem as the workload and the resultant DASD activity changes. When a record is accessed from disk, related records on the same track are also loaded in the cache. Because of the "clustering" nature of disk accesses, a number of subsequent processor requests for data will be serviced by records now in cache rather than requiring another access to the disk. Thus, DASD activity and average DASD response time is reduced. In addition, because the modified LRU algorithm manages the cache contents, tracks from actuators experiencing relatively high access rates tend to remain in the cache longer than tracks experiencing lower access rates. The overall effect is a "dynamic balance" of actuator activity in a string, thus helping to ensure that no single actuator becomes a bottleneck.

DASD Skew—TSO Environment

Figure 2-6 shows the modeled performance effects of DASD skew on a *noncache* 3880 storage director and an eight-actuator DASD string in a TSO/Batch-like environment. The TSO/Batch environment is illustrated but similar results would be typical for IMS, CICS, and other environments. In Figure 2-6, the vertical axis is DASD subsystem response time in milliseconds. The horizontal axis is the accesses per second for a string of 3380 actuators. The curve on the left shows the extreme DASD skew when all of the activity is directed to one actuator. The curve on the far right shows no skew when the activity is evenly distributed over all eight actuators. These two outer curves represent the two extremes for DASD subsystem performance. The effect of skew on performance is to reduce the performance of a DASD string, as measured in accesses per second.



Modeling Data–For Reference Only.

These results are from an analytic model using specific workloads and specific predefined parameters. Individual systems may fall outside these ranges.

Figure 2-6. DASD Skew—TSO Environment

For example, Figure 2-6 shows that with a DASD subsystem response time of 40 milliseconds, a balanced string (activity distributed equally) of eight 3380 actuators can deliver a little over 66 accesses per second. On the other hand, if the workload activity is concentrated to a single device in the string, approximately 21 accesses per second is attained—a reduction of 68%. In contrast, a DASD string with the least skew, in a study of selected production systems, showed a reduction in possible accesses per second of 35%. The highest skew measured showed a 50% reduction in accesses per second. The closer a system is to the equal distribution curve on the right, the better tuned the DASD subsystem is from a data-set-placement standpoint, and the better the performance is from an accesses-per-second standpoint.

DASD Skew—IMS Environment

For an IMS environment, the skew is similar to that of TSO. The workload model showed that at 40 milliseconds response time, a balanced string of 3380 actuators could deliver approximately 70 accesses per second. On the other hand, if the workload is concentrated on a single actuator in the string, fewer than 24 accesses per second would be possible—a reduction of 65%. From a balanced string standpoint, the lowest skew measured showed a reduction in possible accesses per second of 26%; the highest skew measured showed a 46% reduction in accesses per second.

Cache DASD Performance

When comparing the relative performance of cache DASD with noncache DASD, it is necessary to consider the variable of DASD skew. It is also necessary to consider two other variables: hit ratios and read-to-write ratios. These latter variables, although not affecting performance comparisons between similar devices, do alter the effective performance of cache DASD subsystems.

Hit Ratios

An important factor in the performance of a multilevel storage hierarchy is the proportion of processor requests satisfied by finding the data in the fastest storage level compared to those requests satisfied in the slowest storage level. In cache processing, the term "hit" is used to mean that the record referenced by the processor request has a copy in the cache, and a "miss" means that it does not have a copy in the cache.

An algorithm is used to manage data in the cache. The algorithm attempts to ensure that the cache contains only active data just as virtual operating systems ensure that processor storage contains only active pages. In a cache device, the success of this algorithm is indicated by the hit ratio. In general, the higher the hit ratio, the more effectively the cache is being used.

Read-to-Write Ratios

Another important factor in performance is the read-to-write ratio, that is, the number of read requests as compared to the number of write requests directed to cache DASD. Each update of a record in cache is written on disk as part of the input/output operation, because the cache is subject to data loss if a power failure occurs. In this way, the data is protected against loss by ensuring that current copies of the data are always available on disk. A performance penalty is incurred because of the time needed to write on the disk, but data integrity is ensured by this action.

In general, at a given hit ratio, the higher the read-to-write ratio, the lower the average response time for a cache DASD subsystem.

	Transfer Time (ms)
	3.0 MB/sec Channel
Read Hit	3.5
Read Miss	20.6*
Write Hit	23.4
Write Miss	20.6*

* These times assume an average 3380 response time with an 8.0 millisecond seek, an 8.3 millisecond rotational delay, and a single record (4096 bytes) transfer.

Figure 2-7. Performance Characteristics

Figure 2-7 illustrates the transfer times for read and write hits and misses, where:

Read Hit: When the processor accesses a record that is in the cache, it is called a "read hit." A read hit results in the record passing from the cache to the data channel. No access to DASD is required.

Read Miss: When the processor requires a record that is not in the cache, it is called a "read miss." The cache storage director transfers the record from DASD to the data channel and then copies the record from DASD to the cache.

Write Hit: When the processor updates a record that is in the cache, it is called a "write hit" The record is first updated in the cache and then the updated record is copied to DASD to ensure data integrity.

Write Miss: When the processor writes an update to a record that is not in the cache, this is called a "write miss". A normal DASD write operation takes place to record the data on DASD.

Other Considerations

When comparing the maximum capacity of the cache in the Model 13 with the maximum capacity of devices attached to the Model 13, you may wonder how 8 megabytes of cache storage can service up to 20 gigabytes of DASD storage. One factor is that not all data sets on attached devices are opened at the same time. Only active data from opened data sets is in the cache and it is replaced when it is no longer used. Another factor is that once a specific record is read from an open data set, there is a high probability that an adjacent record will be referred to within a short period of time.

Comparing Performances

The following performance information provides a relative comparison between cache and noncache products. The model results that are presented used a fixed set of assumptions in an attempt to ensure a fair and accurate comparison. However, these assumptions may not match any specific user or environment. Therefore, it is important to remember that some of the specified assumptions may not apply for a given system environment. The values shown in the performance curves and tables should not be considered actual obtainable results in a specific environment, but rather as a relative indication of the performance of the product.

The illustrations depict DASD subsystem comparisons using an analytic model with specific workloads and specific predefined parameters. The TSO/Batch processing and IMS workloads used with the model were obtained through studies of representative production systems. A specific system may show significantly different results than these examples.

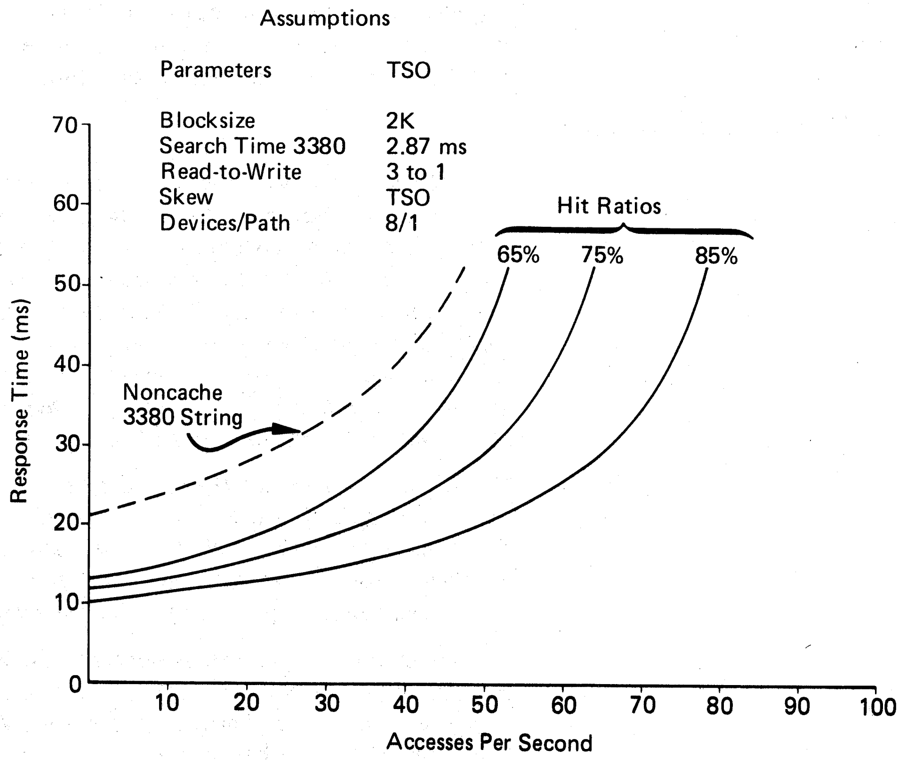
A comparison of cache versus noncache DASD subsystem performance can be made in two ways. One way is to select a constant DASD subsystem response time and compare the number of I/O accesses per second (throughput). The other way is to start with a constant number of I/O accesses per second and compare the differences in DASD subsystem response time. For batch systems, the latter comparison is probably misleading and throughput comparisons should be used. If you can assume that the system is tuned or configured to provide adequate response times for data base and interactive environments, the throughput comparisons are more meaningful. However, if the DASD I/O response time is higher than required for adequate system response, the DASD subsystem response time comparsions should be considered.

In the illustrations that follow, the vertical axis represents the average DASD subsystem response time in milliseconds. The response time is measured from the initiation of the I/O operation (including queues) by the processing system to the receipt of "device end" in the processing system. The horizontal axis represents the number of completed I/O accesses per second, or throughput. Each figure shows the assumed hit ratios of 65, 75 and 85 percent, as indicated by the curves. A performance curve of a noncache disk string is provided for relative comparison.

In the past, many models have shown a uniform access distribution that assumed equal accesses to each device on the string. The intent of showing an "average DASD skew" and an "average read-to-write ratio" is to picture a realistic model.

TSO/Batch Performance

Figure 2-8 illustrates the assumptions for the TSO/Batch model. The analytic model assumed an average skewed access distribution of a TSO/Batch processing workload taken from Figure 2-6. No particular cache size is assumed, only that sufficient cache capacity is available to achieve the designated hit ratios.



Modeling Data–For Reference Only.

These results are from an analytic model using specific workloads and specific predefined parameters. Individual systems may fall outside these ranges.

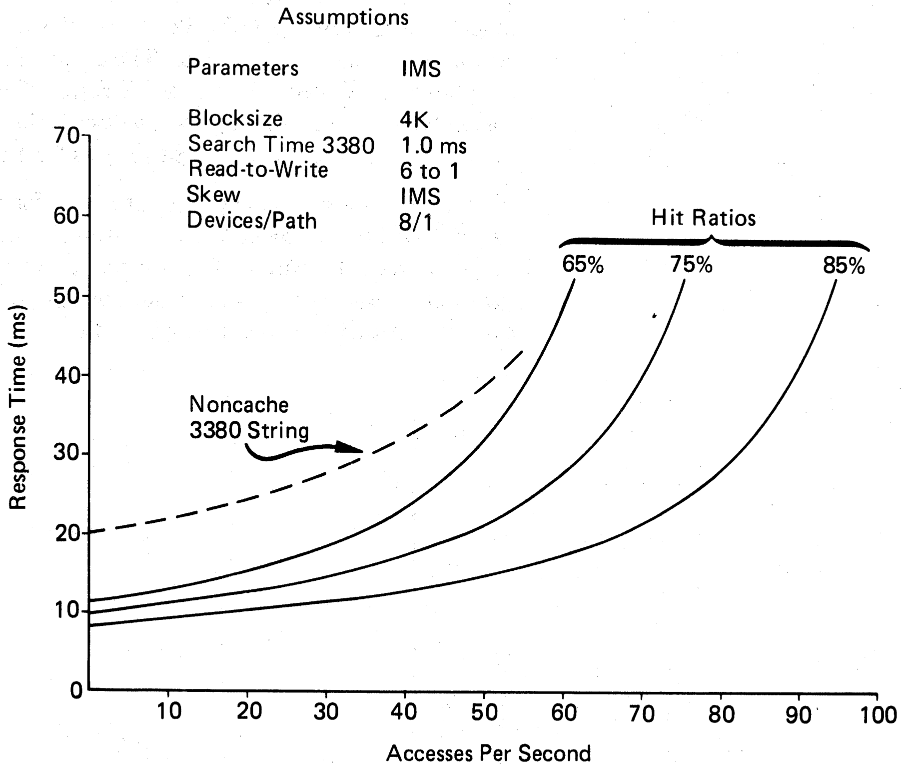
Figure 2-8. TSO/Batch Performance Comparison

The figure also shows the modeled TSO/Batch environment and the performance characteristics of a cache string of eight 3380 actuators as compared to a noncache string of eight 3380 actuators. DASD subsystem response time is shown on the vertical axis and accesses per second are shown on the horizontal axis. An observation that can be made is that at a response time of 25 milliseconds, the noncache 3380 string has 12 accesses per second. Using the same response time of 25 milliseconds and a 75% hit ratio, the 3880 Model 13 string shows a possible 270% gain when compared to the noncache 3380 string.

Another observation that can be made from Figure 2-8 is the difference in average DASD subsystem response times. At 30 accesses per second, the noncache string has an average response time of approximately 34 milliseconds. At 30 accesses per second and a 75% hit ratio, the 3880 Model 13 string has an average response time of approximately 18.5 milliseconds, which is a possible 45% improvement.

IMS Performance

Figure 2-9 illustrates the assumptions for the IMS modeled workloads. The IMS workload assumed an average skewed access distribution taken from data similar to that presented in Figure 2-6. No particular cache size is assumed, only that sufficient cache capacity is available to achieve the hit ratios.



Modeling Data–For Reference Only.

These results are from an analytic model using specific workloads and specific predefined parameters Individual systems may fall outside these ranges.

Figure 2-9. IMS Performance Comparison

The figure shows the modeled performance characteristics of a cache string of eight 3380 actuators as compared to a noncache string of eight 3380 actuators in an IMS environment. DASD subsystem response time is shown on the vertical axis and accesses per second are shown on the horizontal axis.

An observation that can be made is that at a response time of 25 milliseconds, the noncache 3380 string has 20 accesses per second in an IMS environment. Likewise, with a response time of 25 milliseconds and a 75% hit ratio, the 3880 Model 13 string shows a 180% gain in throughput over the noncache 3380 string.

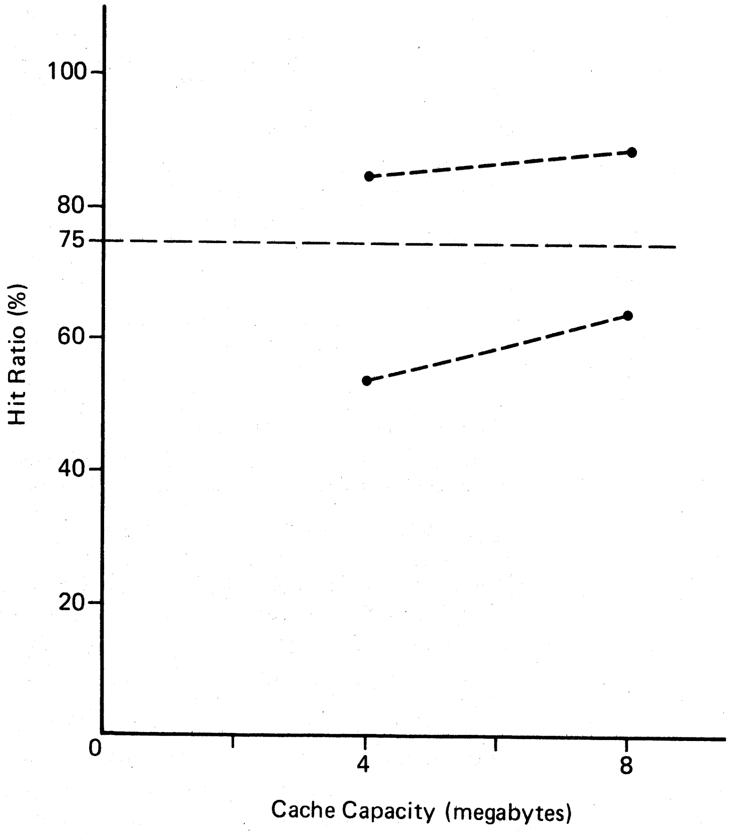
One other observation that can be made from the graph in Figure 2-9 is the difference in average DASD subsystem response times. At 30 accesses per second, the noncache string has an average response time of approximately 28 milliseconds. At the same accesses per second and a 75% hit ratio, the 3880 Model 13 string has an average response time of approximately 14.5 milliseconds, which is a possible 48% reduction in response time.

Cache Size Analysis

The following illustration depicts a simulation of the cache management algorithm of the 3880 Model 13. A specific system may show significantly different results. Individual systems may fall outside these ranges.

A monitor was used to record DASD activity on a group of production systems. This data was then used in a modeling program that simulated the algorithm used in the IBM 3880 Model 13. A range of operating systems and applications were studied. These various DASD activities, when simulated, produced a range of hit ratios. The average hit ratio per sample was calculated and this figure used with the average hit ratio from other systems to produce a combined range of hit ratios.

Figure 2-10 shows the average hit ratios for the MVS systems studied. The two dotted lines connect the highest and lowest average hit ratio points that resulted from the study. The top points represent the maximum average hit ratios at 4- and 8-megabytes; the bottom points represent the minimum average hit ratios of the samples at those same two capacities.



Modeling Data–For Reference Only.

This illustration depicts a simulation of the cache management algorthm of the 3880 Model 13. Individual systems may fall outside these ranges.

Figure 2-10. Cache Capacity versus Hit Ratio

Performance Summary

The actual performance gained using a cache storage director depends upon a number of factors besides the size of the cache. Applications with a low read-to-write ratio will not realize as significant an improvement as will applications with a higher read-to-write ratio. Applications where data is never reused, such as dumping from DASD to tape, will gain no benefit. Applications in which data tends to be randomly used will need a larger cache. Applications involving sequential reading of large data blocks will show less improvement compared to the sequential reading of small data blocks.

Conclusion

Today's data processing environment is characterized by a rapid increase in the number of terminal users. Good terminal response time is needed to increase user productivity and this requires a well-managed DASD subsystem. But frequently DASD bottlenecks are the cause of poor response times. The growing need to share data between processors tends to compound the DASD management problems. Solving these problems, whether by using skilled manpower or by acquiring of extra DASD is often unsatisfactory.

Using cache as a data buffer between DASD and the processor offers attractive solutions to the problem of DASD management. Performance is improved because there is no seek or rotational delay for data that is stored in the cache. The effect of DASD skew on DASD string performance is reduced. Contention for data sets and DASD paths is less because data in the cache is available in a few milliseconds. The IBM 3880 Model 13 offers better DASD performance and improved DASD management to match the demand of the '80s.

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.

Introduction to IBM 3880
Storage Control Model 13
Order No. GA32-0062-0

READER'S COMMENT FORM

This manual is part of a library that serves as a reference source for systems analysts, programmers, and operators of IBM systems. This form may be used to communicate your views about this publication. They will be sent to the author's department for whatever review and action, if any, is deemed appropriate. Comments may be written in your own language; use of English is not required.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation whatever. You may, of course, continue to use the information you supply.

Note: *Copies of IBM publications are not stocked at the location to which this form is addressed. Please direct any requests for copies of publications, or for assistance in using your IBM system, to your IBM representative or the IBM branch office serving your locality.*

	Yes	**No**
• Does the publication meet your needs?	☐	☐
• Did you find the information:		
Accurate?	☐	☐
Easy to read and understand?	☐	☐
Easy to retrieve?	☐	☐
Organized for convenient use?	☐	☐
Legible?	☐	☐
Complete?	☐	☐
Well illustrated?	☐	☐
Written for your technical level?	☐	☐
• How do you use this publication:		
As an introduction to the subject?	☐	
For advanced knowledge of the subject?	☐	
To learn about operating procedures?	☐	
As an instructor in class?	☐	
As a student in class?	☐	
As a reference manual?	☐	

- What is your occupation?

Comments:

If you would like a reply, please give your name and address.

Thank you for your cooperation. No postage stamp necessary if mailed in the U.S.A. (Elsewhere, an IBM office or representative will be happy to forward your comments or you may mail them directly to the address on the back of the title page.)

Reader's Comment Form

Fold and tape **Please Do Not Staple** Fold and tape

Attention: Information Development
Department 61C

NO POSTAGE
NECESSARY IF
MAILED IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 40 ARMONK, NEW YORK

POSTAGE WILL BE PAID BY ADDRESSEE

International Business Machines Corporation
General Products Division
P. O. Box 27155
Tucson, Arizona 85726

Fold and tape **Please Do Not Staple** Fold and tape

IBM®

International Business Machines Corporation
Data Processing Division
1133 Westchester Avenue, White Plains, N.Y. 10604

IBM World Trade Americas/Far East Corporation
Town of Mount Pleasant, Route 9, North Tarrytown, N.Y., U.S.A. 10591

IBM World Trade Europe/Middle East/Africa Corporation
360 Hamilton Avenue, White Plains, N.Y., U.S.A. 10601

GA32-0062-0

Introduction to IBM 3880 Storage Control Model 13 Printed in U.S.A. GA32-0062-0

International Business Machines Corporation
Data Processing Division
1133 Westchester Avenue, White Plains, N.Y. 10604

IBM World Trade Americas/Far East Corporation
Town of Mount Pleasant, Route 9, North Tarrytown, N.Y., U.S.A. 10591

IBM World Trade Europe/Middle East/Africa Corporation
360 Hamilton Avenue, White Plains, N.Y., U.S.A. 10601