StorageTek.

Enterprise Storage Report
for the 1990s

by Fred Moore
Corporate Vice President
Strategic Planning
Storage Technology Corporation

Enterprise Storage Report for the 1990s

Abstract

Data processing has become an increasingly vital function, if not the most vital function, in most businesses today. No longer only a mainframe domain, the data processing enterprise also includes the midrange and workstation platforms, either local or remote. This expanded view of the enterprise has encouraged more and more businesses to take a strategic, long-range view of information management rather than the short-term tactical approaches of the past. This paper will highlight some of the significant aspects of data storage in the enterprise for the 1990s.

ENVIRONMENT - 1990s

- **Storage and storage management are most pressing issues**

- **Networking and connectivity requirements increasing at all levels**

- **DASD subsystem reliability and availability continue to improve**

- **Fault tolerant DASD architecture emerging**

- **Strong acceptance of automated libraries**

- **New focus and mission for magnetic tape technologies**

As the 1990s begin, effective storage management remains possibly the most pressing issue. Poor device utilization and erratic performance are no longer accepted as normal conditions. The cost of ineffective storage management also has received considerable attention as storage costs now exceed the processor costs in mainframe environments.

The definition of the enterprise has moved quickly beyond the world of IBM mainframes to include other mainframes, midrange distributed processors and networks, local and wide area. The need to connect these processing platforms through standard network interfaces and provide access to common storage is increasing rapidly as most users now have mixed-vendor environments (Cray, DEC, IBM, Tandem, etc.) to manage.

The reliability of DASD subsystems continues to improve but even at 99.99 percent availability the only acceptable goal remains 100 percent availability. This trend has encouraged several companies to develop fault-tolerant DASD architectures. Fault-tolerant DASD subsystems provide continuous data availability in the event of any hardware component failure.

The successful introduction of automated tape systems such as StorageTek's 4400 Automated Cartridge System has led to widespread acceptance of automated storage. The 1980's view that library architectures were the least reliable component of a data center is now obsolete. The data processing industry has overcome preconceptions created by various mass storage and rail-type architectures of the past.

The highly successful launch of automated cartridge systems has given new life to tape data storage. Primarily used for low-activity backup, automation has allowed many new applications, not previously considered for tape, to become practical and cost-effective.

ENVIRONMENT - 1990s (Cont.)

- **Automated operations becoming a strategic goal for most large installations**

- **Accelerated new application growth in PC/ workstation segment**

- **Outsourcing slowing**

- **Disk growth rates have moderated to 25-30% annually**

- **Image processing market slowly emerging**

- **IBM SYSPLEX re-focuses mainframe role**

Automated operations is quickly becoming a strategic goal of most large-scale data centers.

For the first time, there is now more storage outside the "glasshouse" mainframe environment than in it. This accelerated growth rate for storage away from the mainframe area will lead to system-managed storage structures, automated tape systems, multi-media libraries and fault-tolerant DASD for the midrange and desktop computing environments.

Outsourcing, a trend that gained considerable visibility in the late 1980s, has lost some of its appeal. Sourcing some or all computer operations, services and development to a source outside the enterprise is intended to save money. Though initial short-term financial gains are possible, many users now are viewing outsourcing as losing control of the most critical component of a business — the information processing function.

Mainframe DASD growth rates have moderated. In the early 1980s, DASD growth rates pushed a 60 percent annual increase in installed gigabytes. As the 1990s unfold, we note that most of the batch-to-online conversions are over. Secondly, data bases are now common, predominant in most organizations, eliminating redundant data. The third reason is that more users are beginning to make more effective use of storage management tools. Finally, many of the new applications that remain to be automated are emerging slowly such as image processing. Even at 25 percent compounded annual growth, the installed base of DASD capacity will double every three years.

It is believed that at the mainframe, midrange and workstation levels, image processing will be the dominant single driving factor for storage demand in the 1990s. This movement, however, has not evolved as quickly as projected due primarily to the lack of an effective enterprisewide image management architecture.

The role of the mainframe in the 1990s was clearly established by IBM's September 5, 1990, announcement of SYSPLEX. This announcement refocused the role of the mainframe in the enterprise as the central server and overseer of the networked enterprise. Mainframe architecture will continually drive many of the standards used for the entire enterprise.

DEMAND TRENDS

	Processor Demand (MIPS)				Storage Demand (GB)
	1960-1969	1970-1979	1980-1989	1990-1999	1990-1999
Mainframes	15%	20%	40%	25%	25-30%
Midrange	- -	35%	50%	40%	40%
Workstation/ Desktop	- -	- -	150%	60%	55-65%

Let's take a look at some of the growth rates that we have seen in the last three decades and will see in the future. This chart projects CPU or processor demand. Notice that processor demand for mainframes in the 1990s is expected to be around 25 percent, measured in MIPS.

We observe 25 percent growth in the mainframe, 40 percent in the midrange, and the workstation growing overall at about 60 percent annually during the 1990s. Today MIPS demand corresponds almost one-to-one with storage growth. Storage management and the ability to access all data objects from all computing platforms will become both a requirement and a major architectural challenge of the 1990s. The vendor that can resolve this problem best will control the enterprise.

ENTERPRISE DATA STORAGE ENVIRONMENT
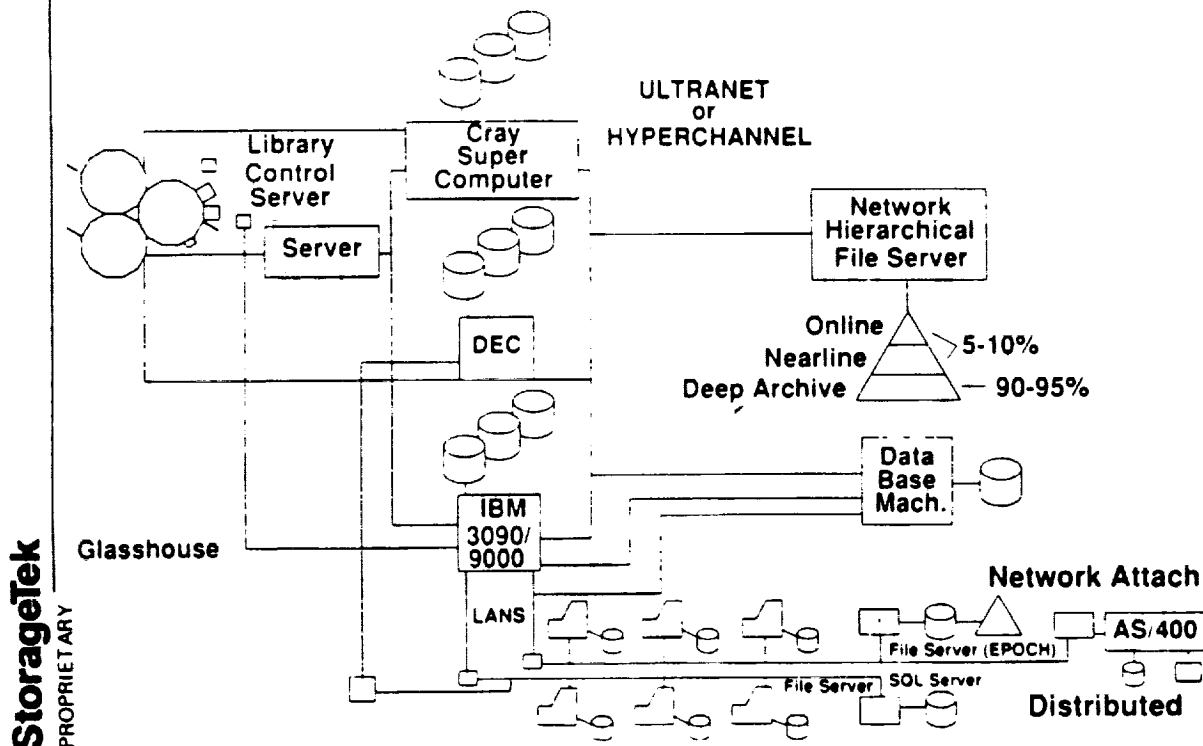
We have had the traditional glasshouse or mainframe view of data processing for a long time. That glasshouse today, dominated by MVS and VM environments, is beginning to share the spotlight with the rapidly emerging midrange and workstation/desktop processing environments. The enterprise now requires management of heterogeneous and complex environments. These three platforms are clearly and distinctly emerging as major areas of processing and storage for the 1990s.

Data processing now becomes distributed at nodes in the enterprise and the objective is to allow transparent access while maintaining the security, integrity and performance of the environment. The role of large systems in the 1990s will become one of management and control for the information enterprise.

We will see the migration of MIPS and storage, and the management issues that go with them, move from mainframe to midrange and desktop. We are not going to be able to limit our views of storage management to MVS and glasshouse and IBM-only for much longer. Storage management solutions must cross those architectural and communication boundaries.

Nearline is a registered trademark of Storage Technology Corp.

VIRTUAL STORAGE HIERARCHY

Average Cost/MB	Performance	Capacity
$4,000	80 ns	4 GB
$1,560	75 μs	16 GB
<$900	1.5 ms	nX GB

Central

Expanded
Storage

Auxiliary
Storage

- Contains operating systems, programs and pages

- Non-shared data, required for HSP1
- Can contain HSP2, data spaces and pages

- Paging devices, SSD and DASD
- Contains HSP2 and data spaces in page datasets

The virtual storage hierarchy of the 1990s will be exploited by ESA/390 architectures and consist of three levels of storage. Central memory, at about $4,000 per megabyte, has an architectural limit of 4 gigabytes and an announced 1 gigabyte limit. Expanded storage has a 75-microsecond access time. The limitation in present ESA/390 architecture is 16 gigabytes of addressable expanded storage, though the announced limit is 8 gigabytes.

Solid-state products are now considerably less than $900 per megabyte and the cost per megabyte is declining quickly. In 1979 when the first SSD was introduced by Storage Technology Corp., the original price was $8,800 per megabyte. We have seen over a 90-percent reduction in pricing on solid-state technology in the 1980 decade. The virtual storage hierarchy contains three levels including auxiliary or paging storage. Careful use of all three technology levels offers the most cost-effective solution to managing the virtual storage hierarchy. It is normally not cost effective for most users to place all performance-critical data in expanded and central storage.

The ES/9000 processor series now permits migrated pages to move directly from expanded storage to the channel subsystem (auxiliary storage paging) improving the synergy and performance between both levels of the virtual storage hierarchy.

LARGE SYSTEMS ARCHITECTURE
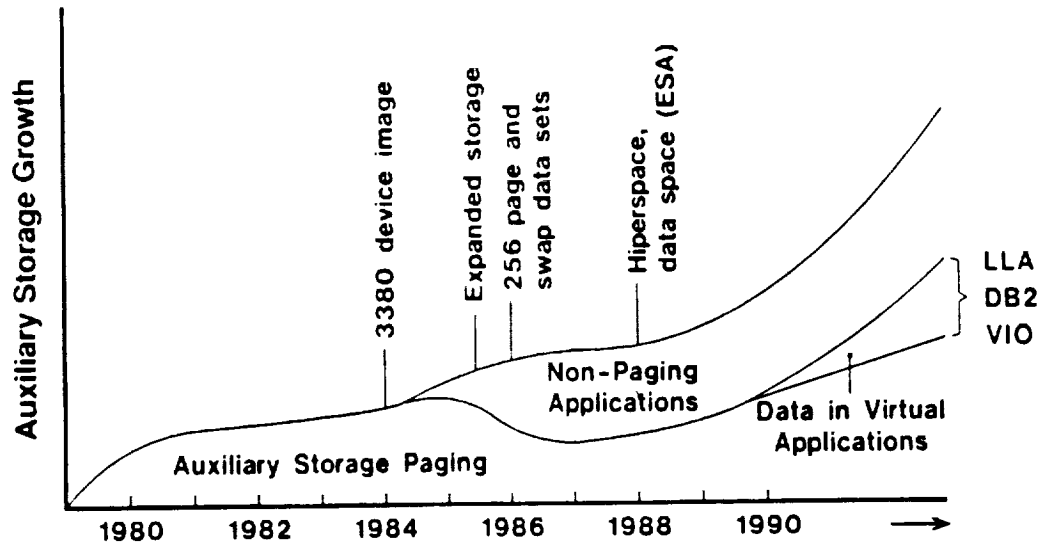## 1993 AND BEYOND
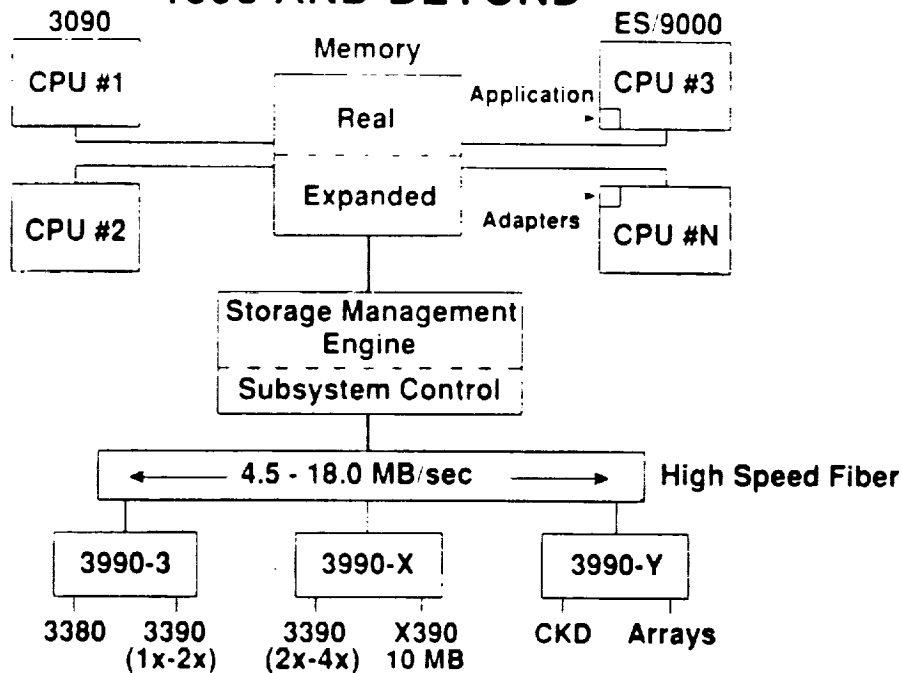
IBM's recent SYSPLEX announcement clearly refocused the role of the mainframe in the 1990s. Let's examine a likely scenario for the 1993 time frame and beyond. This has often been referred to as the post-Summit or Future Systems (FS) architecture. It is expected that up to a limit of 16 (CPU #N) ESA-based processors evolving within a SYSPLEX could be connected in the manner shown.

The continued roll-out of this architecture will include a shared expanded storage capability and application-specific adapters implemented via software, licensed internal code and hardware in varying amounts. Hardware assists for DFSORT announced in the September 5, 1990, IBM announcement are using this concept.

A storage management engine or I/O processor will be a new concept used to off-load from the host processor many of the I/O functions such as parts of IOS (I/O Supervisor), VSAM and DF (Data Facility) functions. The storage management engine will attach peripheral devices as we know them today (SSD, DASD, tape, printers and terminals) via the channel subsystem. Attachment of ESCON (Enterprise System Connectivity) serial fiber channels will be preferred though parallel bus and tag channels will need to attach via ESCON converters. The point-to-point limit of ESCON channel transfer rates is 18 MB/sec.

ESCON is a trademark of IBM Corporation

DIRECTION OF AUXILIARY STORAGE
RAM BASED ARCHITECTURES

Auxiliary Storage Growth

3380 device image

Expanded storage
256 page and
swap data sets

Hiperspace,
data space (ESA)

LLA
DB2
VIO

Non-Paging
Applications

Data in Virtual
Applications

Auxiliary Storage Paging

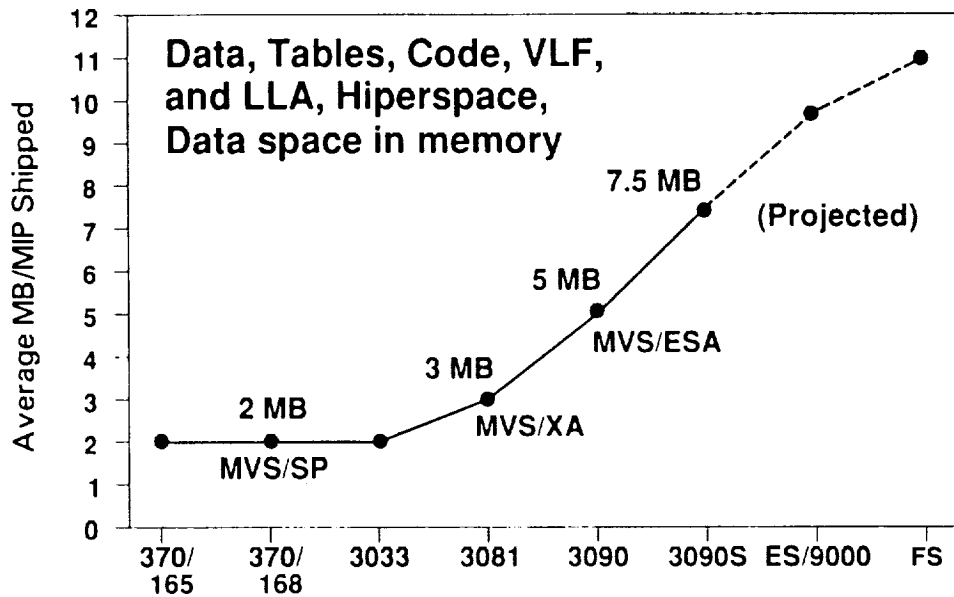1980   1982   1984   1986   1988   1990

By examining RAM-based solid-state products, you will notice their use from 1979 through 1984 was exclusively for paging. When expanded storage appeared in 1985, solid-state devices became viewed as a high-performance disk, and non-paging data such as load libraries, catalogs and indexes were placed on solid-state devices. The 1990s will see the MVS/ESA and VM/ESA hiperspace and data space applications drive up data in virtual requirements and force users to seriously consider using SSD as a cost-effective complement to real and expanded storage. ESA/390 drivers of virtual storage consumption, called methods of I/O avoidance by some, will include linear VSAM, Virtual Lookaside Facility (VLF), hiperspace catalog, hiperspace buffers and DB2. MVS/XA systems previously using 400 to 500 megabytes of auxillary storage will soon identify requirements exceeding 1 gigabyte or more after migrating to ESA.

DRAM MEMORY DENSITY PROJECTIONS

Technology (in microns)	DRAM Technology	Development Start	Introduction Date	Peak Output
1.8	256 Kbit	1977	1982-83	1988
1.2	1 Mbit	1980	1985-86	1991
0.8	4 Mbit	1983	1988-89	1994
0.6	16 Mbit	1986	1991-92	1997
0.36	64 Mbit	1989	1994-95	2000
0.25	256 Mbit	1992	1997-98	2003
0.15	1 Gbit	1995	2001-01	2006

StorageTek PROPRIETARY

Unlike rotating DASD, RAM-based architectures command a very price-elastic market. If the price decreases, the demand increases. You cannot necessarily stimulate the demand for DASD or tape by changing the price. The industry-standard DRAM chip has moved from 1 megabit to 4 megabits. Note that the 4-megabit chip has been under development since 1983. As DRAM densities increase, price decreases along with the physical space required to store information. Thus much higher capacity DRAM storage devices will appear occupying smaller footprints. This trend should continue until the point where DRAM-based storage devices will occupy a large portion of the storage hierarchy currently belonging to rotating and cached DASD.
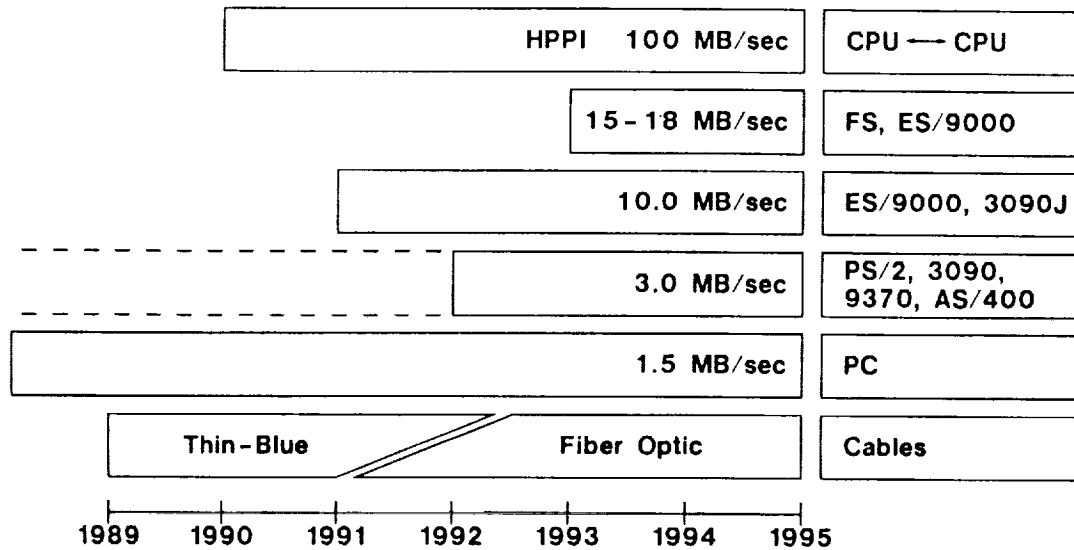
MEMORY/MIPS SHIPPED

Data, Tables, Code, VLF,
and LLA, Hiperspace,
Data space in memory

Average MB/MIP Shipped

7.5 MB

(Projected)

5 MB

MVS/ESA

3 MB

2 MB

MVS/XA

MVS/SP

Historical tracking of the ratio of memory installed per MIPS installed indicates the trend increasing
sharply with MVS/XA (31-bit addressing) and MVS/ESA, effectively 44-bit addressing. The MVS/ESA
capability to place data, and program load libraries and other objects into memory, will continue to
drive the ratio upward, requiring more and larger RAM-based storage solutions. The announcement
of VM/ESA and DOS/VSE/ESA implementing hiperspace and data space concepts into these oper-
ating systems will further encourage virtual storage growth. The growth rate grows sharply until
shared expanded storage arrives late in the Summit (i.e., future systems) then moderates slightly.
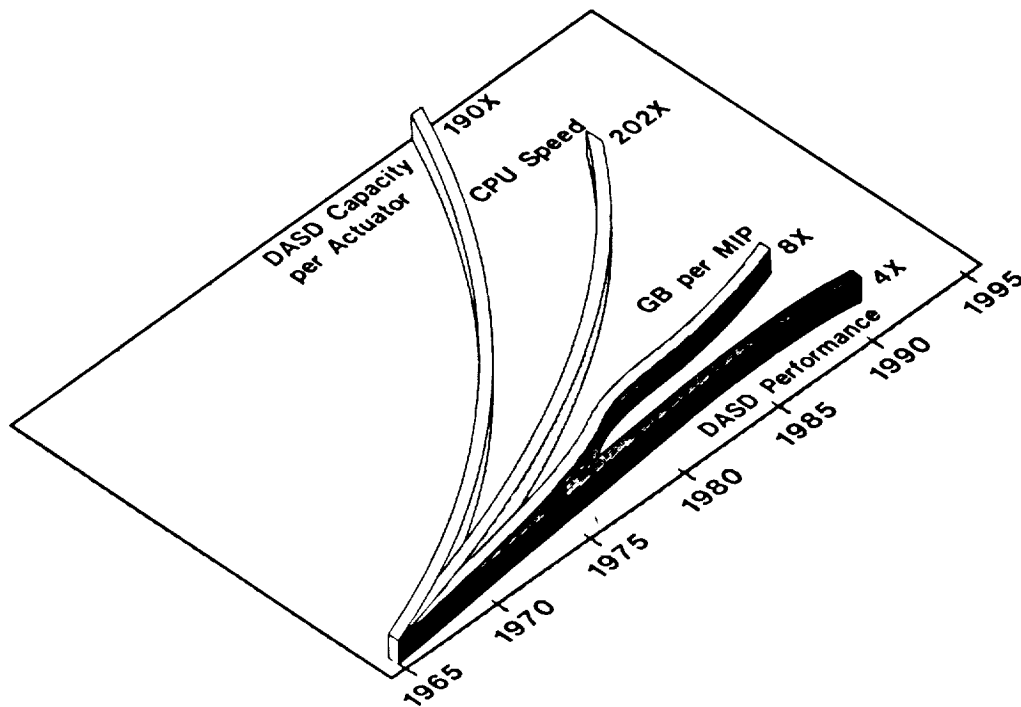
CHANNEL TRANSFER RATE
GROWTH PROJECTION

Rate	Device
HPPI 100 MB/sec	CPU — CPU
15 – 18 MB/sec	FS, ES/9000
10.0 MB/sec	ES/9000, 3090J
3.0 MB/sec	PS/2, 3090, 9370, AS/400
1.5 MB/sec	PC
Thin – Blue Fiber Optic	Cables

1989   1990   1991   1992   1993   1994   1995

StorageTek
PROPRIETARY

The data transfer rate capabilities at all levels of computing will increase much faster in the first half of the 1990s than they did in the previous 15 years. ESCON channels now offer up to 10 MB/sec data transfer rates. Device exploitation of ESCON channels at 10 MB/sec will come more slowly. The 3990-3 DASD Storage Control will be the first device to exploit ESCON channels at ESCON speeds. Up to 18 MB/sec on the FS series is likely by 1995. The gray and thin blue cables will begin to give way to serial fiber channels providing increased distance (up to 9 km initially) in the early 1990s. The HPPI (High Performance Parallel Interface) channels used by large-scale CPUs for scientific processing will offer attachment of specialized storage devices in the mid-1990s, likely RAM-based, providing high-performance solid-state storage arrays.

LARGE SYSTEMS TRENDS

DASD Capacity per Actuator   190X

CPU Speed   202X

GB per MIP   8X

DASD Performance   4X

1965 1970 1975 1980 1985 1990 1995

In the last 25 years, the processing power of large computers has increased at a much faster rate than the performance capabilities of the I/O subsystem. Since the introduction of System/360 in 1965, we have seen the capacity of a disk actuator increase 190 times. Processor performance has increased over 200 times, but the performance of a disk actuator has improved only 4 times. This divergence of processor speed and the I/O subsystem performance has been the subject of considerable interest, particularly in the 1980s. During this time, we have seen the introduction of a number of technology developments to help bridge the gap. These enhancements include solid-state disk, cached control units, dual port, quad port, actuator level buffers, tape buffers and expanded storage. Despite these advances, the performance gap between processors and I/O subsystems continues to diverge. More solutions will emerge, predominately based on DRAM technology, to place data closer to the processor and remove the performance delays of mechanical devices. The ES/9000 processor announcement by IBM is a good example of this continually diverging trend — processor speed (MIPS) nearly doubled while the speed of the storage subsystem remained unchanged.

TRENDS – AREAL DENSITY

StorageTek
PROPRIETARY
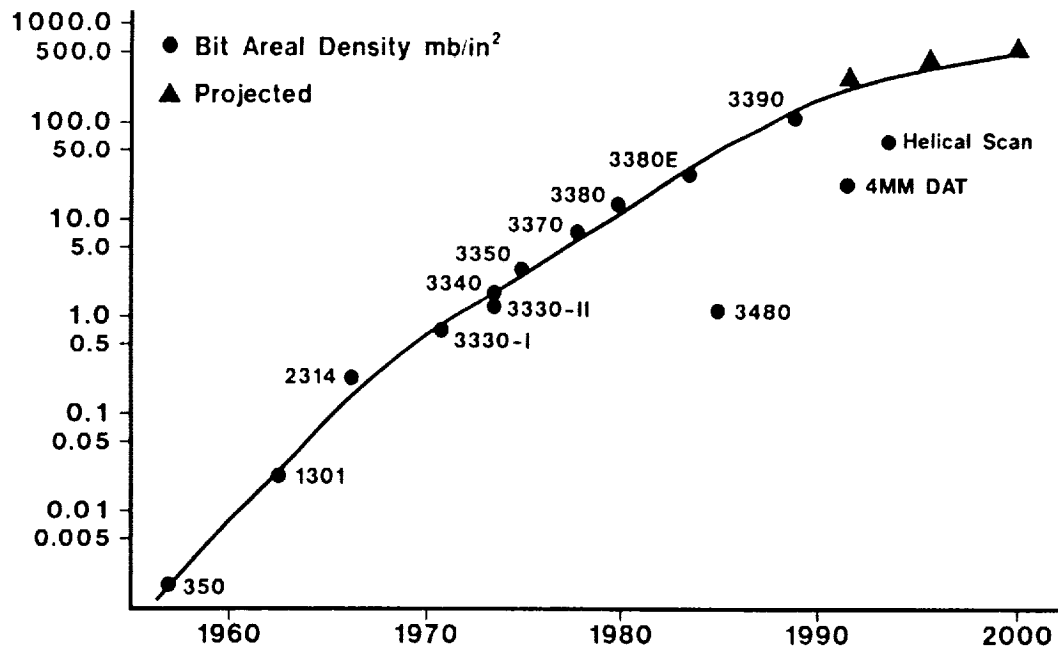
Magnetic recording technology has seen few of the anticipated limitations in the areal density (mega-bits per square inch) that were expected to occur in the early 1980s. Areal density increases in DASD, once not expected to exceed 30 megabits per square inch, are now over 60 megabits per square inch in 3390-type disk drives and are expected to increase to over 500 megabits per square inch by the end of the 1990s. Laboratory developments of 1 gigabit per square inch have been demonstrated. As the areal density of magnetic recording continues to increase rapidly, the future role of optical disks in the large systems storage hierarchy becomes more questionable.

The helical scan tape format, recording data tracks on a tranverse rather than parallel to the edge of the tape, offers areal densities in the range of 30-50 million bits per square inch and very high data transfer rates. Helical scan technology should enter the hierarchy at the deep archive level and co-exist with 3480 chromium dioxide (Cro2) tape format throughout the remainder of the 1990 decade.

OPTICAL
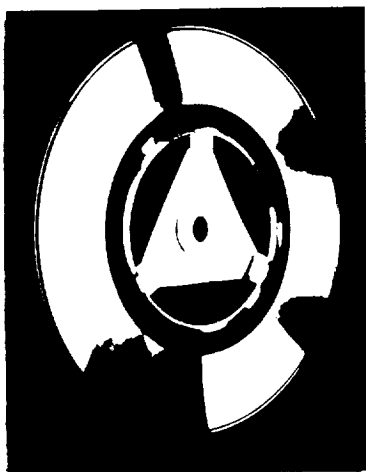
WHEN

MAGNETIC/OPTICAL STORAGE
5 1/4" OPTICAL PARAMETRICS vs DASD

OPTICAL HDA	1990	1995	2000
Capacity (GB)	1.0	1.5	4.0
Transfer rate (MB/sec)	0.7	1.2	2.5
Cost ($/MB) (OEM)	3-5	1.30	0.50
DASD HDA			
Capacity (GB)	1.5	2.5	5.0
Transfer rate (MB/sec)	1.9	4.0	6.0
Cost ($/MB) (OEM)	2.0	0.60	0.20

9-90 FORUM90220J08

Optical disk storage has been hailed as the low-cost mass-storage technology of choice for years. In reality, optical storage has yet to fulfill expectations. Issues such as standardization, throughput, data transfer rates and uncertainty of shelf life (data retention) remain. The areal density of magnetic disk storage is increasing rapidly, while optical disk areal density has remained relatively the same for the past six years. The write-once, read many times (WORM) drives are common in the midrange and desktop markets, but will struggle to find a mainframe niche. Magneto-optic or erasable optical disks offer the large systems market the most benefit, but may face a stiff challenge from large-capacity DASD array storage solutions for the online, large-capacity storage market and advanced automatic cartridge systems for even larger capacity and less costly deep archive storage. Unlike WORM optical, 5.25" magneto optical has the support of formal standards by the International Standards Organization.

TAPE LIBRARIES
## MAINSTREAM APPLICATIONS

CURRENT APPLICATIONS

- Tape Management
- DASD Management
  - SMS
  - DF/HSM, DMS/OS
  - DASD savings
- Job scheduling and rerun
  - Improved batch performance
- Software development

- Automated recovery
  - Online data bases
  - Mission-critical data
- Report management/paper/fiche
- Electronic archive
  - Campus
  - Remote vault
- Automated operations
  - Unattended → Lights out
- High speed search applications

ADVANCED APPLICATIONS

- Anticipatory staging
  - Random access data
  - Data set scheduling

- Network storage
- Deep archive with helical formats
- Scientific data

StorageTek

Automation has opened many new horizons for data storage. Far beyond simply automating what currently exists on tape today, automation has become a primary ingredient in cost-effective storage management and is enabling the promise of systems managed storage to be fulfilled. Several new applications listed above have become areas of opportunity providing cost savings and improved quality of operations beyond prior capabilities. The concept of electronic archiving has been given an increased focus with the announcement of ESCON channel architecture complementing traditional channel extension methods. A form of image storage, report management, has been enhanced with a number of software products that allow computer output microfiche and printed data to be stored on a tape library, viewed at a computer terminal, and printed or sent to fiche only if needed. This new area of library exploitation greatly reduces distribution, copy and filing costs while improving the security aspects associated with printed storage. Applications with much promise for automated tape libraries include anticipatory staging of data and deep archive storage for long term-data storage.
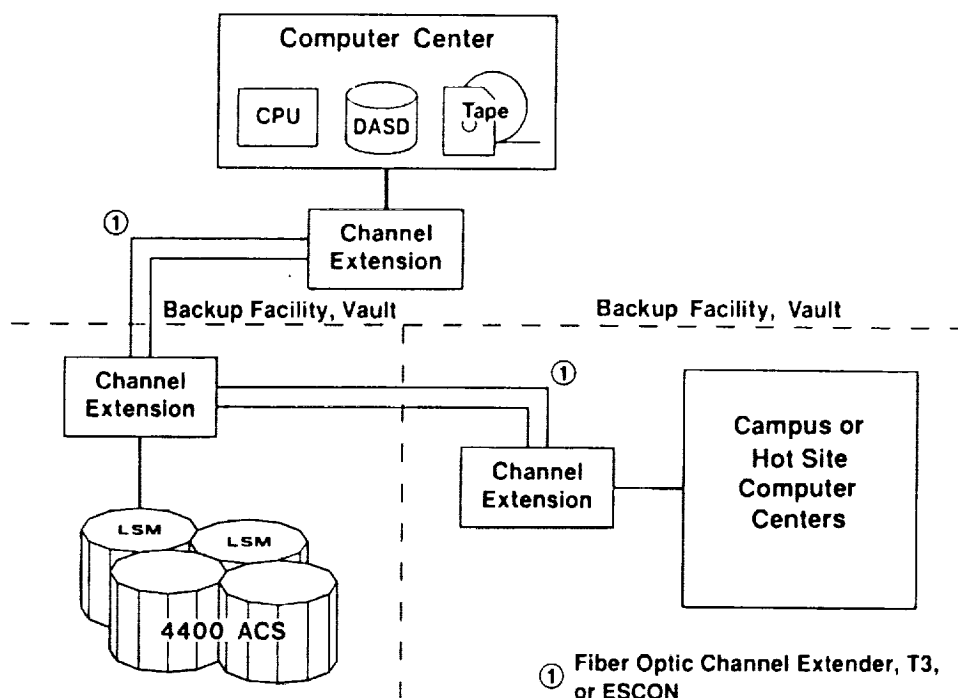
AUTOMATED OPERATIONS

Present ⟶

Manual Processes

Automated Processes

| Phase | 1 | 2 | 3 | N |

Time ⟶
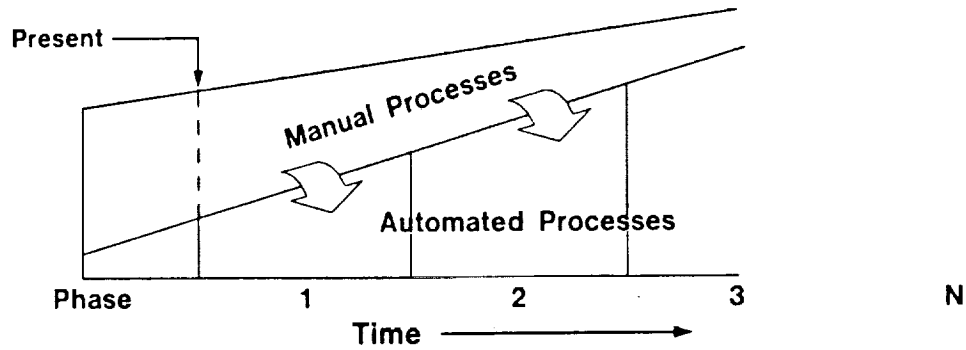
Console Message	Operator	Suppress Auto Reply	Open Architecture
Recovery	Manual	Auto Restart	D/B Recovery Appl. Recovery
SMS	Storage Pools	Space Mgmt. DF/SMS	Performance Management
TMS	Physical Placement	Scratch Retention	ACS 4400
Output Distribution	Burster Trimmer	Roll Feed Bar Code	Online Viewing Selective Print

As automated operations becomes a strategic goal for many data processing users, solutions are appearing which are making companies more competitive, more productive and more profitable. Automated operations is usually fully implemented in stages and will evolve to include expert systems solutions to resolve some of the complex, enterprisewide information management issues. The primary reason for automated operations is improved quality of the data processing organization.

In addition to automated operations, business resumption or disaster recovery planning has become a strategic goal for many users. Workshops on these areas of advanced data center operations are available on a worldwide basis from Storage Technology Corp.

ELECTRONIC ARCHIVING

Computer Center

CPU | DASD | Tape

Channel Extension

① Backup Facility, Vault ——— Backup Facility, Vault

Channel Extension

Channel Extension

Campus or Hot Site Computer Centers

LSM | LSM

4400 ACS

① Fiber Optic Channel Extender, T3, or ESCON

The rapid acceptance of library storage products such as StorageTek's 4400 Automated Cartridge System (ACS) has provided a means for computer users to archive data electronically to a secure, remote location such as a vault or warehouse. Today, the use of fiber optic channel extensions provides 3 megabyte per second device attachment. Products such as the 4400 ACS can be located at distances well beyond the four walls of the computer center by using fiber optic channel extenders, T3, or ESCON communication lines. The "data vault" provides backup of critical data in a safe location and also can link into a hot site or campus computer facility for a quick recovery in case of a disaster. This trend will expand in the 1990s as the value of information to the corporate enterprise becomes increasingly more important.

Progress of Storage Technologies
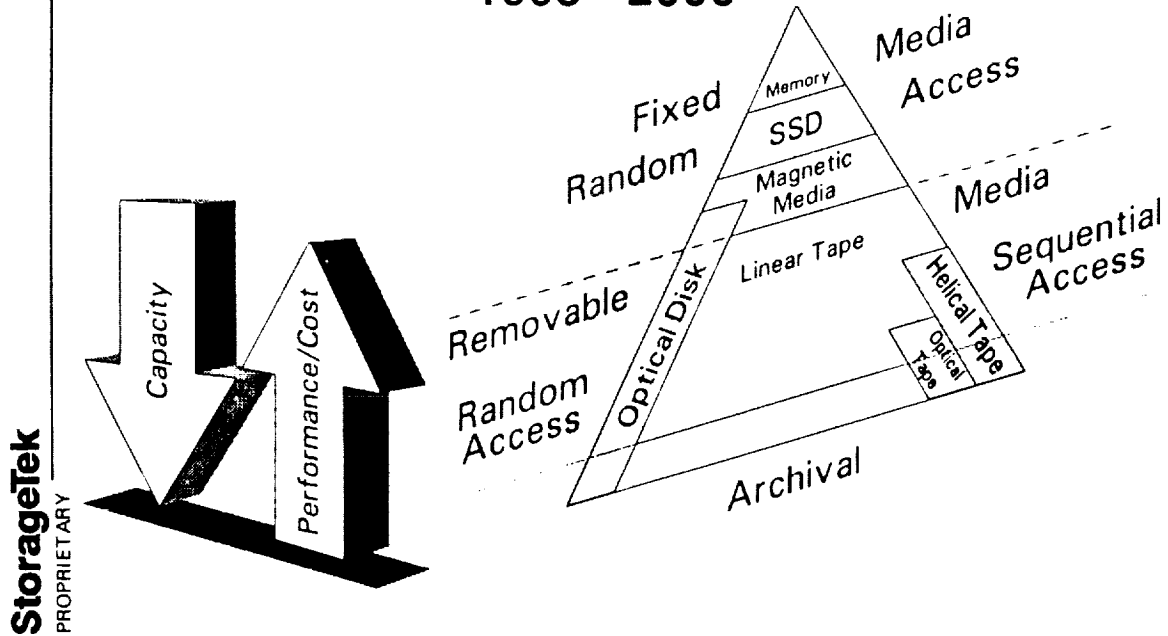

StorageTek
PROPRIETARY
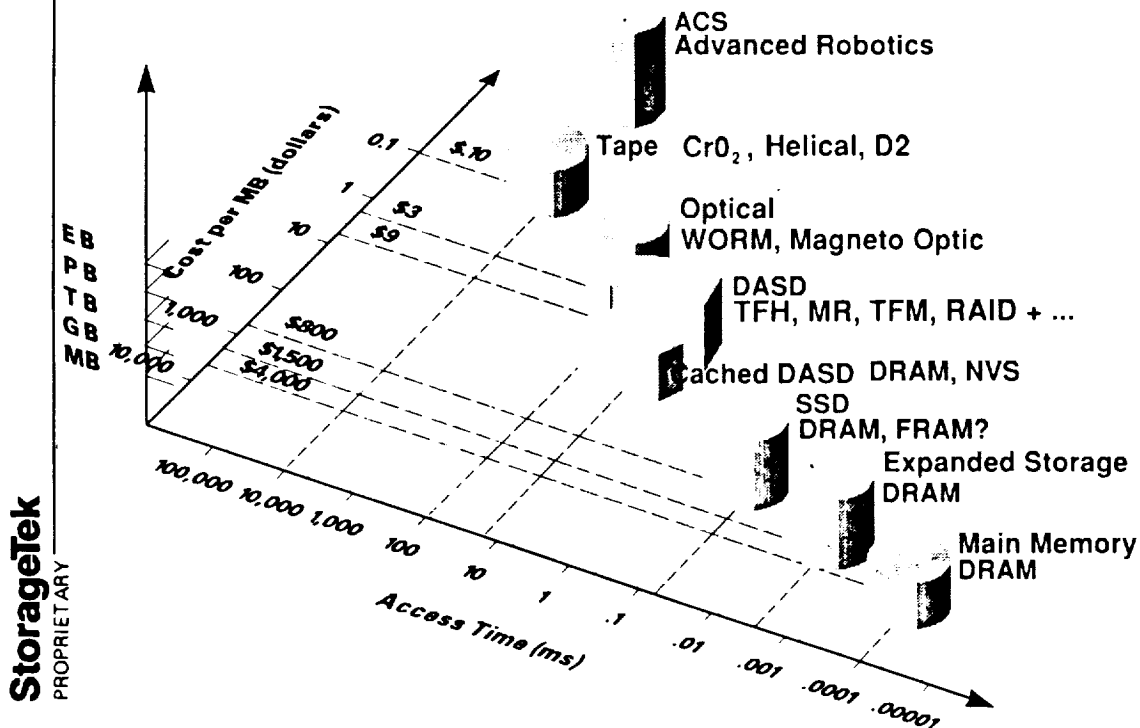
Let's take a look also at a few of the technologies that have merged to help resolve some of the challenges in the 1990s. The 1980s was the the decade of technology. The 1990s will be the decade of how we exploit that technology effectively. In the 1980 timeframe, we clearly remember vendor and customer discussions regarding DASD. Issues centered on such things as the diameter of the disk platter. How thick is the platter lubrication? How high does the read/write head fly? The answers to those questions sometimes influenced buying decisions. Today in DASD acquisitions the issues are gigabytes per square foot; I/Os per second; availability; cost per gigabyte. The size of the platter really doesn't have to make a difference, but gigabytes per square foot should.

Notice that main memory presently is priced at about $4,000 per megabyte with an access time measured in nanoseconds. Expanded storage has a 75 microsecond access time and is priced at about $1,500 per megabyte. There remains a major performance and price gap between the memory technologies (DRAM) and moving or rotating technologies.

Optical storage and tape provide two interesting comparisons. The areal density of optical storage has witnessed insignificant improvement in the last six years. During this time, magnetic storage has significantly increased in areal density. Interestingly, the IBM Image Plus system, using a write once optical storage device, is priced around $2 to $3 per megabyte. Tape libraries, including compression/ compaction, may realize costs as low as 10 cents per megabyte.

New technologies will evolve to fill this gap such as ferro-electric RAM (FRAM) devices. These are non-volatile RAMs and still under development though expected to be affordable and commercially available in the 1994 time frame.

STORAGE HIERARCHY
1995 - 2000

The hierarchy of storage technologies, not devices, in the last half of the 1990s is shown. This hierarchy may be broken down into fixed-media and removable-media segments. Fixed-media storage will consist of RAM and rotating DASD. Optical disk, once viewed as the heir apparent to the removable-media segment, has given way to automated tape systems that are faster and less costly. Further advances in the capacity of 3480 cartridge capacity to 1 gigabyte levels and the increased usage of helical scan formats make magnetic tape (along with magnetic disk) the key technologies of the 1990s.
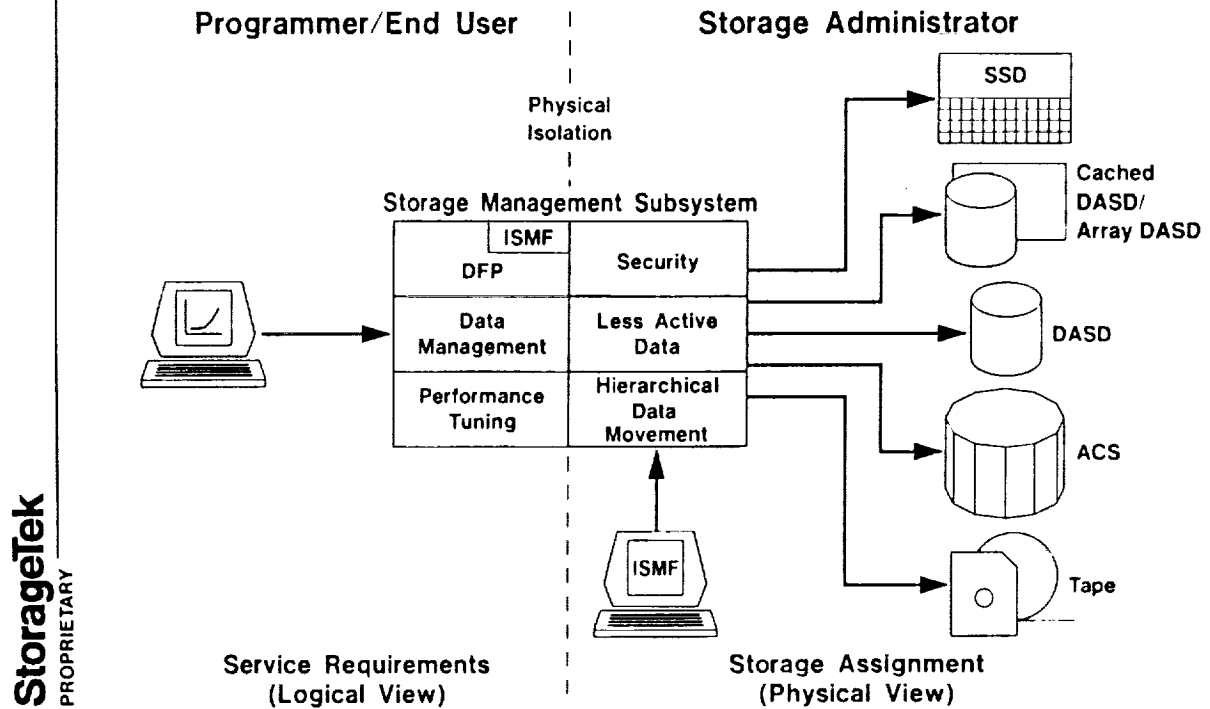
TRENDS – STORAGE MANAGEMENT EVOLUTION
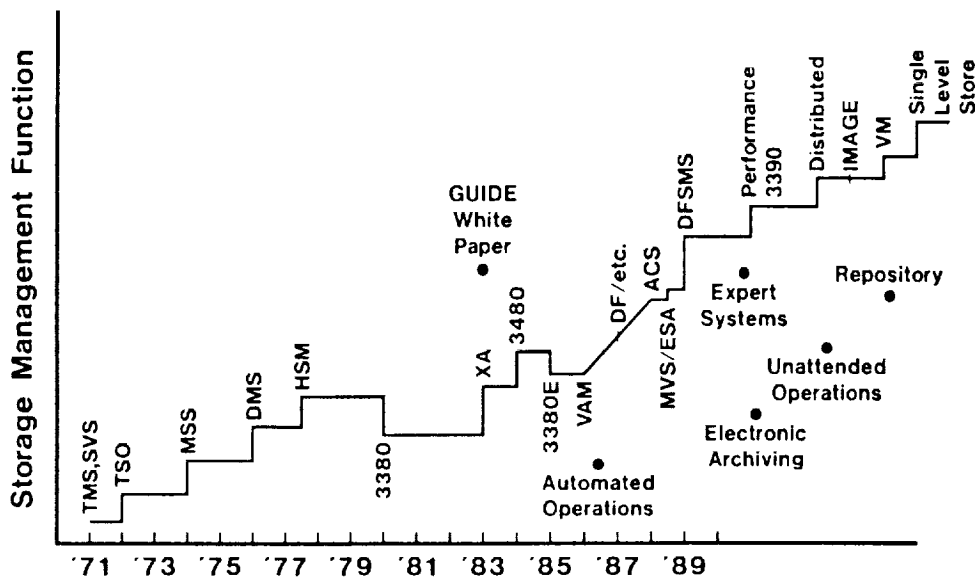
The total cost of storage devices now accounts for more than one half of the total hardware expenses of the typical large data center today. In the early 1970s, storage management meant tape management. With the introduction of the DMS/OS and HSM storage management products in the late 1970s, storage management expanded its scope to include space management for disks. Since that time, storage management had been relegated to improving various facets of space management until the announcement of DFSMS in February 1988. This platform should gradually evolve to include dynamic performance tuning, storage management for distributed processing nodes, networks, workstations, a DFSMS equivalent for VM and development of a repository to identify objects across all computing platforms in the corporate enterprise.

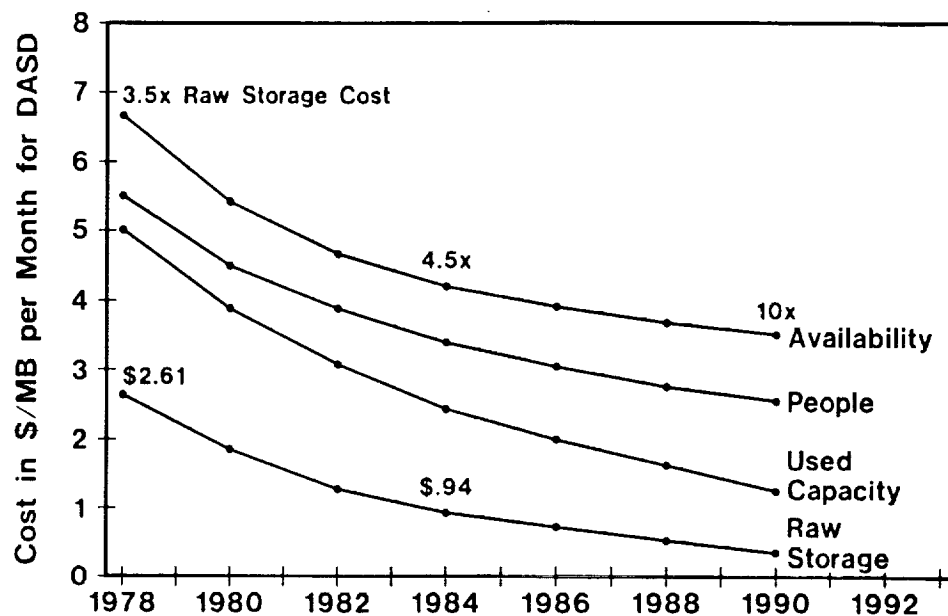DFSMS is a trademark of IBM Corporation

SYSTEMS MANAGED STORAGE

Programmer/End User   **Storage Administrator**

Physical
Isolation

Storage Management Subsystem

ISMF	
DFP	Security
Data Management	Less Active Data
Performance Tuning	Hierarchical Data Movement

SSD

Cached DASD/ Array DASD

DASD

ACS

ISMF

Tape

Service Requirements
(Logical View)

Storage Assignment
(Physical View)

Systems managed storage is a concept that allows an operating system to perform many of the human-intensive processes involved with space management, performance tuning, availability management and true hierarchical storage management supporting all tiers of storage. Though in its infancy, systems managed storage must efficiently evolve these processes to provide the platform for single-level storage in the last half of the 1990s and achieve an environment that allows "true systems managed storage." DFSMS is one of several products that make up systems managed storage. Presently the DFSMS product provides no performance tuning capabilities or movement of data vertically throughout the storage hierarchy to optimize performance or space management.
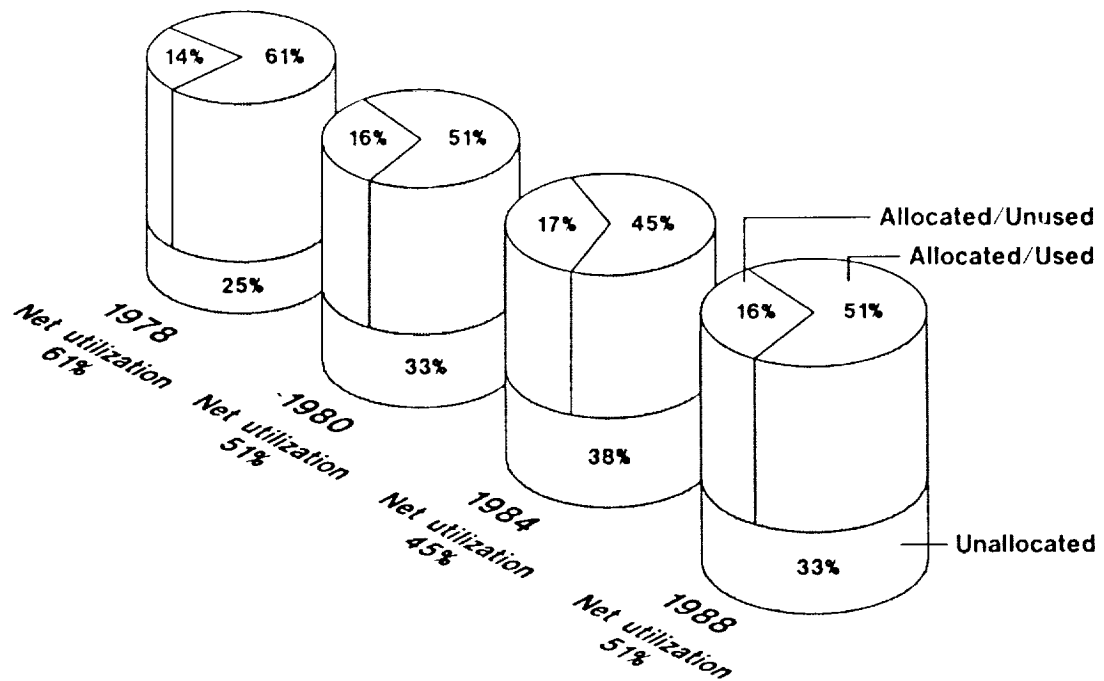
COST OF USABLE, MANAGED STORAGE

3.5x Raw Storage Cost

4.5x

10x
Availability

People

Used
Capacity

Raw
Storage

$2.61

$.94

Cost in $/MB per Month for DASD

1978 1980 1982 1984 1986 1988 1990 1992

Source: IBM Systems Journal, 1989

The total cost associated with managing storage has declined since 1978 on a per-megabyte basis. The cost of raw (live) data, plus the costs of unused capacity, people costs to manage the storage system and additional costs of availability such as backing up data, have increased several times since 1978 on a per-megabyte basis. The cost of managing DASD, effectively or ineffectively, has become a primary concern in very large (terabyte plus) data centers and clearly is necessitating the movement toward much improved storage management facilities. It is estimated that the total cost of managing disk storage in the 1990s will be as much as 10 times greater than the cost of actual data stored on DASD. New storage solutions, such as advanced fault tolerant DASD architectures, will dramatically improve these trends.

DASD USAGE TRENDS

14%   61%
25%
1978
Net utilization
61%

16%   51%
33%
1980
Net utilization
51%

17%   45%
38%
1984
Net utilization
45%

16%   51%
33%
1988
Net utilization
51%

— Allocated/Unused
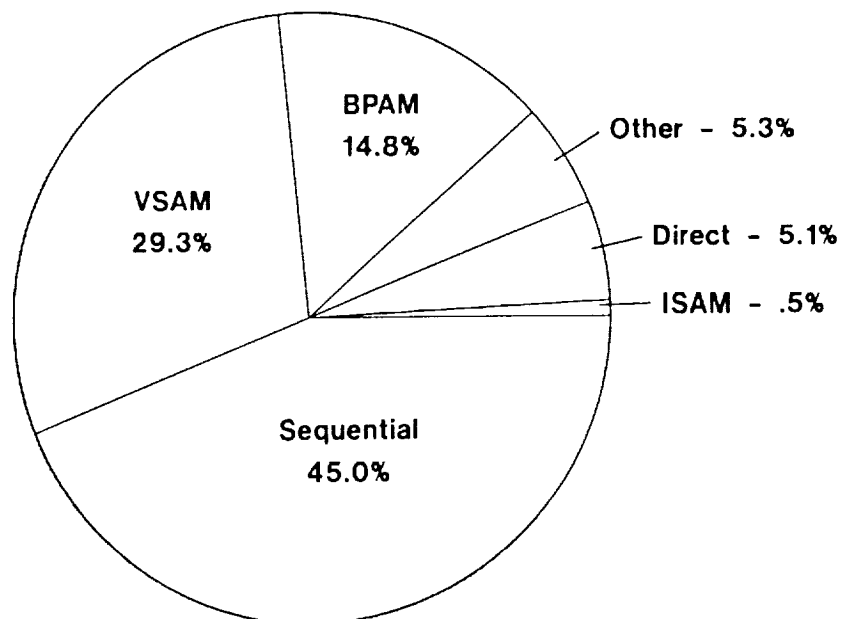— Allocated/Used

— Unallocated

Formal surveys on DASD usage have been conducted since 1978. The net utilization, or the amount of real data on the average device decreased from 61 percent in the 1978 survey to a low of 45 percent in 1984. The 1988 survey indicated an overall increase in net utilization to 51 percent. This survey included single-, double- and triple-capacity 3380-type devices. The single- and double-capacity devices actually increased in utilization; however, the triple-capacity devices continued the downward trend. Utilization of the single- and double-capacity devices increased largely due to the higher percentage of caching permitting increased space allocation on cached DASD. Space utilization figures for 3390-class devices are not presently available though the lack of widespread DFSMS usage to optimize data allocation on 3390-type DASD may initially inhibit more effective utilization.

PERCENT OF DATASETS
BY DATASET ORGANIZATION

VSAM
8.0%

BPAM
16.0%

Other – 2.5%
Direct – 1.4%
ISAM – .1%
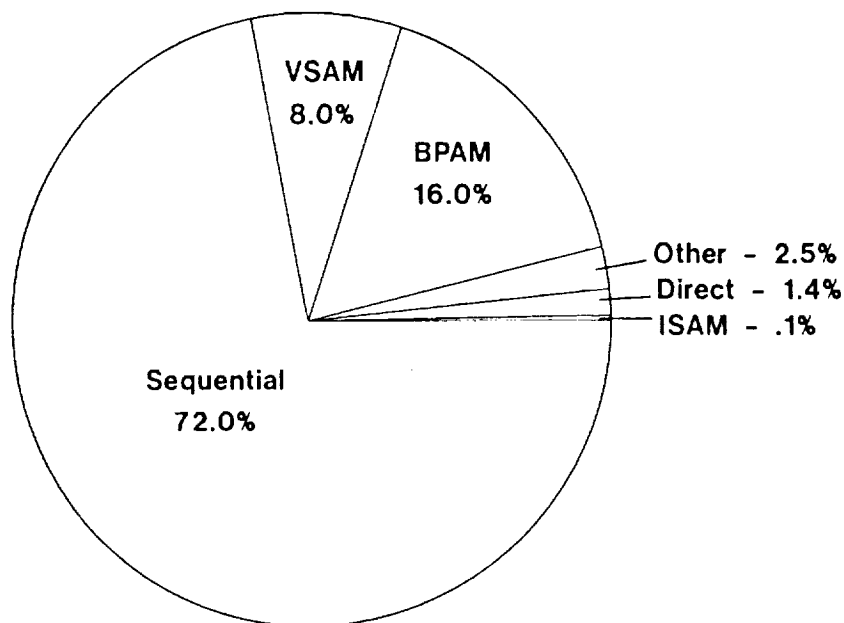
Sequential
72.0%

As a follow-up to the previous chart on DASD space allocation, the percent of DASD data sets by data set organization reveals a correspondingly high percentage of sequential data sets. VSAM and SAM-E (sequential) data set organizations are strategic while partitioned data sets will fold into VSAM format as ESA/390 evolves. Other data sets such as graphics access methods, direct access files and even ISAM will exist as they are today. This profile again reflects the results of extensive tape-to-disk migration activities in the 1980s.
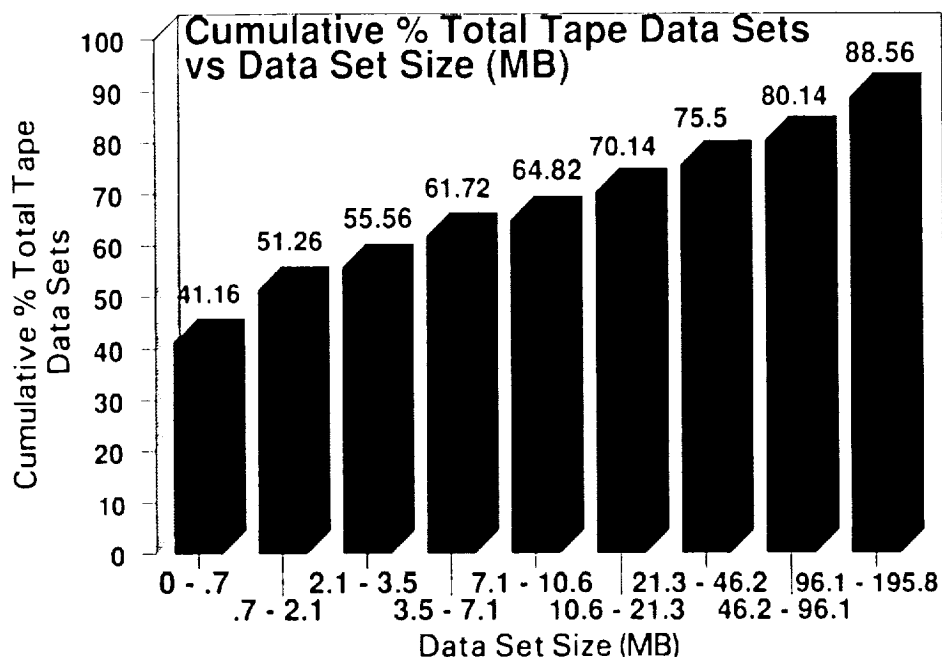
PERCENT OF SPACE
BY DATASET ORGANIZATION

BPAM
14.8%

Other – 5.3%

VSAM
29.3%

Direct – 5.1%

ISAM – .5%

Sequential
45.0%

Storage management has increasingly focused on DASD in the last few years, primarily due to the criticality of data residing on DASD. At the end of 1988, sequential data had grown to include 45 percent of allocated disk space. This was primarily a result of the large number of tape-data-set-to-disk-conversions in the early 1980s, occurring for the lack of any successful automated tape library system available to mainframe users. Tape data sets requiring rather quick or frequent access could not often withstand erratic human tape mount times.
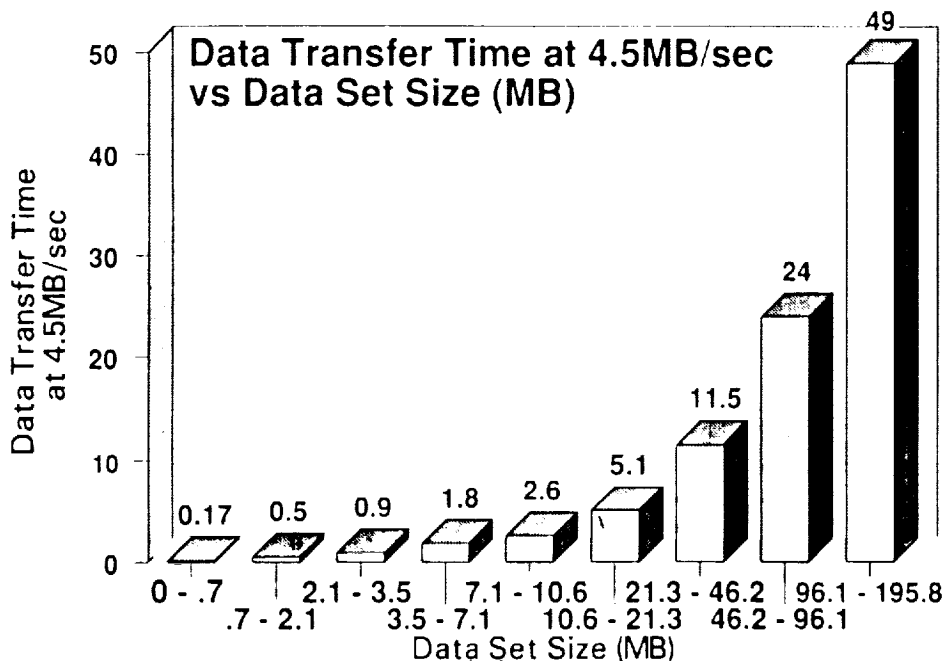
This DASD profile, as a result of the 1980s strategy of "put it all on disk," has become a significant cost savings target for automated cartridge systems in the 1990s.

TAPE COMPRESSION/COMPACTION
TAPE DATA SET PROFILES

Cumulative % Total Tape Data Sets vs Data Set Size (MB)

Bar chart values:
- 0 - .7: 41.16
- .7 - 2.1: 51.26
- 2.1 - 3.5: 55.56
- 3.5 - 7.1: 61.72
- 7.1 - 10.6: 64.82
- 10.6 - 21.3: 70.14
- 21.3 - 46.2: 75.5
- 46.2 - 96.1: 80.14
- 96.1 - 195.8: 88.56

Y-axis: Cumulative % Total Tape Data Sets (0 to 100)
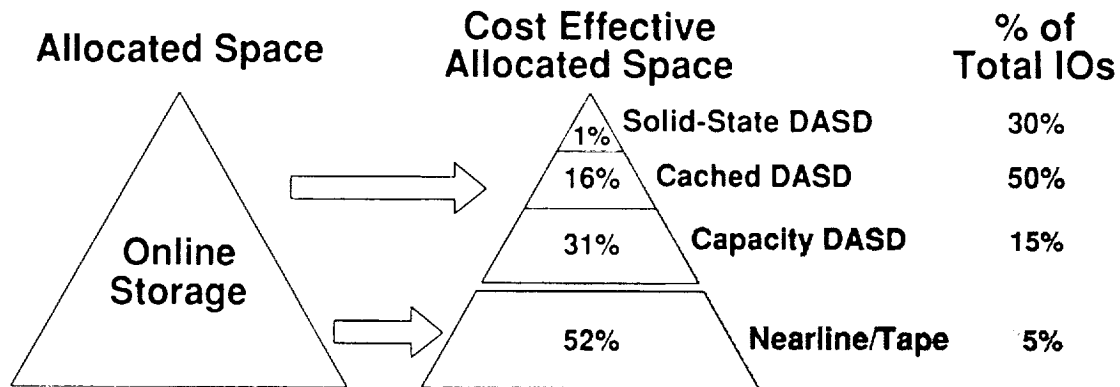X-axis: Data Set Size (MB)

A survey of tape data set sizes indicates that the vast majority of tape data sets are very small. 70.14 percent of the tape data sets in the survey are under 21.3 megabytes in size. Even with the tape-to-disk conversions of the 1980s, many small tape data sets remained on tape and became obvious targets for automation. While 11.44 percent of the tape data sets occupy over one 200MB 3480 tape cartridge, the majority of these data sets are backup applications using processor data-compression functions.

TAPE COMPRESSION/COMPACTION
TAPE DATA SET PROFILES

Data Transfer Time at 4.5MB/sec vs Data Set Size (MB)

Y-axis: Data Transfer Time at 4.5MB/sec (0, 10, 20, 30, 40, 50)

Bar values: 0.17, 0.5, 0.9, 1.8, 2.6, 5.1, 11.5, 24, 49

X-axis: Data Set Size (MB)
0 - .7 | .7 - 2.1 | 2.1 - 3.5 | 3.5 - 7.1 | 7.1 - 10.6 | 10.6 - 21.3 | 21.3 - 46.2 | 46.2 - 96.1 | 96.1 - 195.8

StorageTek
PROPRIETARY

The tape data set size survey is equated to data transfer time for each classification of data set on a 4.5 MB/sec channel using bus and tag or even ESCON. The 21.3 MB tape data set takes approximately 5.1 seconds to process once mounted. Assuming a 2-to-1 data-compression factor, a total time of 2.55 seconds would be used to transfer the 21.3 MB data set. This minimal gain from compression/compaction is often lost by the variability in manual tape mounts resulting in minimal, if any, throughput benefit for 70 percent of the tape data sets. By the early 1990s, nearly everything written to tape will be compressed. Like 7-track, 9-track, NRZI and GCR in the past, tape data compression will be another format for tape data recording.

COST-EFFECTIVE STORAGE PROFILES

Allocated Space	Cost Effective Allocated Space	% of Total IOs
	1% Solid-State DASD	30%
	16% Cached DASD	50%
	31% Capacity DASD	15%
Online Storage	52% Nearline/Tape	5%

1980s
- Technology decisions
- Get the right data in the right place

1990s
- Business decisions
- Get the right data in the right place at the right time

Using advanced data storage placement methodologies such as CPIO (Cost Per I/O) analysis, it is possible to determine the most cost-effective location in the storage hierarchy. Using the size, performance and storage cost per I/O to determine optimal data set placement, studies indicate that online storage users today may not be cost effectively utilizing online storage. Typically, 1 percent of the space allocated in online storage generates 30 percent of the total online IOs. This small but highly active group of data sets is most cost effectively located on solid-state disk. At the other end of the spectrum, a little over half of the data allocated on DASD today is more cost effective on automated cartridge systems or manual tape. Storage management in the 1980s stressed getting the right data in the right place; storage management for the 1990s will stress getting the right data in the right place at the right time.

CPIO is a proprietary software tool from Storage Technology Corp.

notes