The Essential Guide to Computer Data Storage

From Floppy-to-DVD

- The complete decision-maker's guide to choosing—and using—the right data storage technologies
- Personal and small business storage: hard drives, DVD, CD, Zip®, Jaz®, and more
- Enterprise storage: SAN, NAS, Fibre Channel, RAID, and beyond
- New Web-based storage alternatives
- Start-to-finish backup and disaster recovery tips

Dr. ANDREI KHURSHUDOV

MICROSOFT CORP. EXHIBIT 1026
Page 1 of 27

Library of Congress Cataloging-in-Publication Data available

Editorial/Production Supervision: *MetroVoice Publishing Services*
Acquisitions Editor: *Mike Meehan*
Editorial Assistant: *Linda Ramagnano*
Buyer: *Maura Zaldivar*
Art Director: *Gail Cocker-Bogusz*
Interior Series Design: *Meg Van Arsdale*
Cover Design: *Bruce Kenselaar*
Cover Design Direction: *Jerry Votta*
Project Coordinator: *Anne Trowbridge*

© 2001 Prentice Hall PTR
Prentice-Hall, Inc.
Upper Saddle River, NJ 07458

Prentice Hall books are widely used by corporations and government agencies for training, marketing, and resale.

The publisher offers discounts on this book when ordered in bulk quantities.
For more information, contact Corporate Sales Department, phone: 800-382-3419;
fax: 201-236-7141; e-mail: corpsales@prenhall.com
Or write: Prentice Hall PTR
Corporate Sales Department
One Lake Street
Upper Saddle River, NJ 07458

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

All product names mentioned herein are the trademarks or registered trademarks of their respective owners.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

ISBN 0-13-092739-2

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Pearson Education Asia Pte. Ltd.

The Modern Computer

From the application point of view, there are up to four large groups of computer systems available today that have different needs for processing power, memory, size and weight, network computing capabilities, and, what is most important for this book, the size and characteristics of storage.

Mobile computer systems (laptop, notebook, hand-held, etc.) are built for high mobility, word processing, multimedia applications, network computing, Internet access, and other forms of online communication. These computers require lightweight, large capacity, and power-efficient storage. Data access can be comparably slow than that for the *server* class (discussed later) since its primary applications are not as time-intensive. This storage has to be tolerant to frequent "on" and "off" operations. The devices working with mobile computers must support plug-and-play features and be as easy to use as possible.

Desktop computers for home and office are meant for network computing, Internet access, email, word processing, spreadsheets, finance software, graphics, and multimedia applications. The storage, therefore, has to be large, relatively fast, but, most importantly, inexpensive. The types of interfaces used most often are ATA/IDE (the dominant), ATAPI, parallel, serial, USB, and, more often today, IEEE 1394 FireWire.

High-end professional *workstations* are used for real-time digital signal and image processing, video production, heavy-duty computations, 3D graphics and special effects, animation, and so on. The type of storage needed here is based on such interfaces as Ultra SCSI, Ultra2 SCSI, Ultra160 SCSI, Ultra160+ SCSI, and Fibre Channel (FC). Apparently, price is not a major concern here. Performance is the main quality required.

Server computer systems are designed for multiuser multiapplication environments such as the Internet or Intranet. These computers deal with intensive data and image processing, video editing, data mining, online transaction processing (OLTP), large databases, and so on. Storage for server computers needs to have high capacity, fast performance, and great reliability. These systems are also quite expensive and rely on the same high-performance interface types as the workstations.

These four groups of computer systems are different in speed, memory, storage capacity, size, and so on, but inside, they all have the same structure and same key elements. Figure 1–4 shows a simplified schematic of a computer.

The main elements of a computer are:

- CPU (central processing unit), or just processor. In human terms this is the equivalent of the brain.
- Memory or RAM (random access memory) is the working area where programs are run and data is processed. Fast, expensive, and volatile equiva-

MICROSOFT CORP. EXHIBIT 1026
Page 3 of 27

lent of human memory with the exception that when you wake up (computer power-on) after sleep (power-off), you still remember something from yesterday. Computer RAM doesn't.

- System bus (Memory bus, IDE bus, PCI bus, etc.). Performs a function similar to linking different organs in the human body. Works like the neurons and moves commands and information from one part of a PC to another. Works like the heart when it moves blood (power) to different devices.
- Storage. Can be compared to a long-term memory, which keeps information one doesn't use quite often and that requires effort and time to access.
- Output and input ports are equivalents of our ears, eyes, speech, writing, and other ways of communication and information exchange.

This book concentrates on *storage*, discusses the *interface*, and just touches on the subject of *memory*.

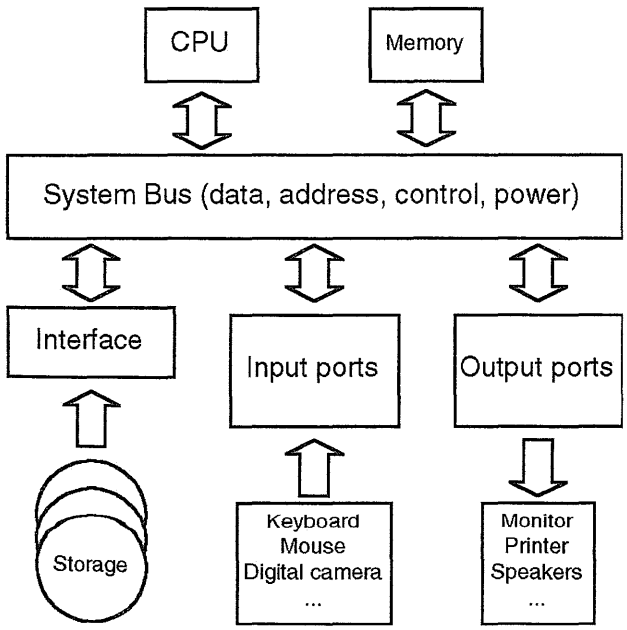


Figure 1–4
A simplified schematic of a computer.

MICROSOFT CORP. EXHIBIT 1026
Page 4 of 27

Memory and Storage

First of all, we need to define the terms *memory* and *storage*, which will be used extensively in this book.

Computer memory is the place where a computer can store data and programs and execute them as needed.

PC memory (see the memory module in Figure 1–5) is random access memory (RAM), which enables the computer's CPU (central processing unit) to access required instructions and data at a high speed. The reason for the CPU to keep all important information in the memory is that this information can be accessed much faster than from computer storage, such as hard disk drive, CD-ROM, and so on.

Figure 1–6 shows a kind of hierarchy of computer memory and storage. Apparently, there is always a compromise between the capacity, data access rate, and

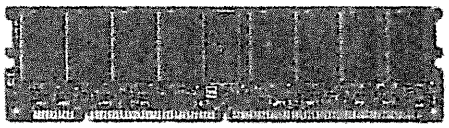


Figure 1–5
168-pin DIMM random access memory.

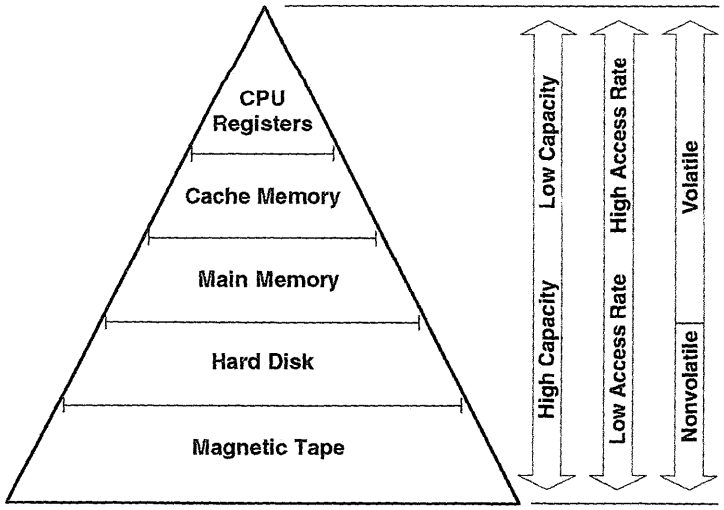


Figure 1–6
The hierarchy of computer memory and storage systems.

MICROSOFT CORP. EXHIBIT 1026
Page 5 of 27

volatility. Also, cost is another important factor; the closer we come to "perfect" memory (large, fast, and nonvolatile), the higher the cost.

Semiconductor memory, for instance, is very fast (access time of 50 to 100 ns compared to 10 ms for hard disk drive) and can be made nonvolatile (i.e., Flash memory). But the cost of this memory is orders of magnitude higher (per unit storage) than that of hard disk or tape drive. This is why different types of memory and storage are used for different tasks.

People sometimes use the terms *storage* and *memory* concurrently, since both of them are used to store computer information. But they are quite different, since the hard disk drive or CD-ROM are not the place where a computer can store data and programs and execute them when necessary, as was defined above. Hard disk drive, CD-ROM, floppy drive, and so on, are the places where a computer can store data and programs *only*. When the program needs to be executed, it has to be transferred from the *storage* into the *memory*.

Storage is meant for less frequently used data or data that could be accessed at a relatively slow speed (still very fast). Memory is the place where the computer keeps its most needed and frequently used data, which can be accessed at much faster speeds. An analogy with the file cabinet (storage) and the work table (memory) is sometimes used to illustrate this difference. One keeps the materials he uses frequently right in front of him on the desk (memory), while all other materials are kept in the file cabinet (storage). Clearly, if one will keep too much data on the desk, it will become increasingly difficult to locate the needed item, and the access speed will suffer. Also, a very large desk will be too expensive. This is why computer memory is typically much smaller than computer storage.

Another critical difference between storage and memory is that storage is always *nonvolatile*, while memory is mostly *volatile* (see Flash memory for nonvolatile memory).

This means the following: when the power is turned off, the nonvolatile memory (hard disk drive, tapes, CD, etc.) retains all the information stored while the volatile memory (DRAM, SRAM, etc.) loses all the information stored.

This book concentrates on *nonvolatile data storage*, or just *storage*.

Binary System

This part of the introduction is needed for understanding how the data is represented in the computer when it is being manipulated and stored.

Everywhere, except for digital computer-related operations, the main system of mathematical notation today is the decimal system, which is a base-10 system. As in other number systems, the position of a symbol in a base-10 number denotes the value of that symbol in terms of exponential values of the base. That is, in the decimal sys-

MICROSOFT CORP. EXHIBIT 1026
Page 6 of 27

Table 1–1 Binary Addition and Multiplication

+	0	1	×	0	1
0	0	1	0	0	0
1	1	10	1	0	1

Addition, subtraction, and multiplication are done in a fashion similar to that of the decimal system:

```
 101101        1011010        1011
+100111        -110101      x 101
1010100         100101       1011
                            0000
                           1011
                           110111
```

The simplicity of the binary operations is one of the reasons why the binary system is used in digital computers. Another reason is discussed next.

Bits, Bytes, and Beyond

Since only two digits, or states (on and off, 1 and 0) are involved, the binary system is used in computers, which are digital devices. The "on" position corresponds to a 1, and the "off" position to a 0 (in fact, this assignment can be opposite with the same result).

In magnetic storage devices (hard disk drives, floppy drives, tape drives, etc.), a magnetization of small areas of the media is used to represent binary information: a change in direction of magnetization between two of these areas stands for 1, and no change stands for 0. Flip-flops—electronic devices that can only carry two distinct voltages at their outputs and that can be switched from one state to the other state—can also be used to represent binary numbers; the two voltage levels correspond to the two digits. Optical and magneto-optical storage devices also use two distinct levels of light reflectance or polarization to represent 0 or 1.

Bit is an abbreviation for *binary digit*—the smallest unit of information in a digital world. A bit is represented by the numbers 1 and 0, which correspond to the states on and off, true and false, or yes and no. Bits are the building blocks for all information processing that goes on in digital electronics and computers. The term bit was introduced by John Tukey, an American statistician and early computer scientist. He first used the term in 1946, as a shortened form of the term binary digit. Bits are usually combined into larger units called *bytes*.

MICROSOFT CORP. EXHIBIT 1026
Page 7 of 27

A byte, in computer science, is a unit of information built from bits. One byte equals 8 bits. The values that a byte can take on range between 00000000 (0 in decimal notation) and 11111111 (255 in decimal notation). This means that a byte can represent 2^8 or 256 possible states (0–255).

Bytes are combined into groups called *words*. The size of the words used by a computer's central processing unit (CPU) depends on the bit-processing ability of the CPU. A 32-bit processor, for example, can use words that are up to four bytes long (32 bits). The term byte was first used in 1956 by German-born American computer scientist Werner Buchholz to prevent confusion with the word *bit*. He described a byte as a group of bits used to encode a character. The eight-bit byte was created that year and was soon adopted by the computer industry as a standard.

Computers are often classified by the number of bits or bytes they can process at one time, as well as by the number of bits or bytes used to represent addresses in their main memory (RAM). A computer graphic is described by the number of bits used to represent *pixels* (short for picture elements), the smallest identifiable parts of an image.

In monochrome images, each pixel is made up of one bit. In 256-color and grayscale images, each pixel is made up of one byte (eight bits). In true color images, each pixel is made up of at least 24 bits. The particular sequence of bits in a byte encodes a unit of information such as a keyboard character. One byte typically represents a single character such as a number, letter, or symbol.

Regular users or professional programmers use computers and software to basically combine bytes in complex ways and create meaningful data in the form of text files or binary files (files that contain data to be processed and interpreted by a computer). Bits and bytes are the basis for representing all meaningful information and programs in computers.

The number of bits used by a computer's CPU for addressing information represents one measure of a computer's speed and power. Computers today often use 16, 32, or 64 bits.

Bytes are also the major unit for measuring quantities of data or data storage capacity. Data quantity is commonly measured in *kilobytes* (about 1,000 bytes), *megabytes* (about 1 million bytes), or *gigabytes* (about 1 billion bytes). A regular floppy disk normally holds 1.44 megabytes of data, which equates to approximately 1,440,000 keyboard characters. At this storage capacity, a single disk can hold a document approximately 700 pages long, with 2,000 characters per page.

Just a few years ago, megabytes represented the largest quantity of data available. Nowadays, gigabytes (GB) are already a common thing, and it is almost time to learn the next measure of data, such as the terabyte.

Table 1–2 Some Capacity Conversions from Binary to Decimal Values

Term	Abbreviation	Binary power	Binary value	Decimal value
Kilobyte	KB	2^{10}	1,024	1,000
Megabyte	MB	2^{20}	1,048,576	1,000,000
Gigabyte	GB	2^{30}	1,073,741,824	1,000,000,000
Terabyte	TB	2^{40}	1,099,511,627,776	1,000,000,000,000

Storage Capacity Definitions

We note that the terminology "1 KB" actually means 1,024 bytes (which is 2^{10}), not 1,000 bytes. The difference is actually quite small and, in most cases, unimportant. Table 1–2 summarizes some capacity conversions from binary to decimal values.

In a hard disk drive, real capacity can be calculated using the following formula (the terms used will be defined in Chapter 3):

```
Capacity = Heads × Sectors × Cylinders × 512 (bytes per sector)
```

Disk drive manufacturers use decimal values of capacity. For example, an 8.4 GB hard disk drive with 16,383 cylinders, 16 heads, and 63 sectors has a storage capacity equal to 8,455,200,768 bytes. Therefore, the drives we buy are slightly larger in capacity than is advertised.

DIGITAL AND ANALOG SIGNALS . . . . . .

Analog versus Digital

Most signals in nature are *analog* and their characteristics vary continuously with time. The sounds we hear, the light we see, the temperature we sense, all of them represent a continuous signal—a data stream with a specific value at any given instance of time. The movements of the trees in the forest or the waves in the ocean are registered by our sensors as a continuous analog signal.

Human vision, hearing, and speech are analog too. When we speak, for example, we generate continuous sound waves that travel in the air by creating pressure fluctuations. When these waves are received by the listener's ear, they cause the eardrum to

MICROSOFT CORP. EXHIBIT 1026
Page 9 of 27

Track

A concentric set of magnetic bits on the disk is called a track. Each track is (usually) divided into 512-byte sectors (see Figure 3–32).

Sector

A part of each track defined with magnetic marking and an ID number. Sectors have a sector header and an error correction code (ECC). In modern drives, sectors are numbered sequentially (see Figure 3–32).

Cylinder

A group of tracks with the same radius is called a cylinder (see Figure 3–32).

Data Addressing

There are two methods for a drive's data addressing: CHS (cylinder-head-sector) and LBA (logical block address). CHS is used in most IDE drives, while LBA is used in SCSI and enhanced IDE drives (see Figure 3–32).

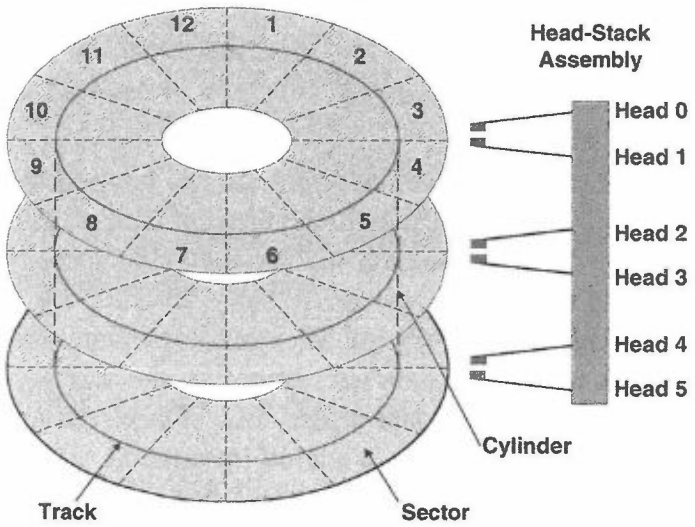


Figure 3–32
Logical design of the hard disk drive.

CHS locates data by simply specifying the cylinder (radius), head (platter and its side), and sector (angular position). LBA assigns each sector of the drive a sequential number, which is simpler.

If you look into your BIOS, you will find listed the number of cylinders, heads, and sectors for each drive you have. The BIOS makes it possible for a computer to boot itself. On PCs, the BIOS contains all the code required to control the disk drives, keyboard, display, serial communications, and some other functions. The BIOS is typically found on a ROM (read-only memory) chip, which comes with the computer.

Modern operating systems access data using LBA directly without the help of the BIOS. This helps to reduce incompatibilities.

Cache or Buffer

To improve performance and increase data rates, HDD utilizes a small amount of fast solid-state memory to store the most frequently used data. This memory is called cache, or buffer. It bridges data from the disk to the interface, since they operate at different data transfer rates. Retrieving data from the cache is much faster than from the disk (microseconds compared to milliseconds). In modern drives, cache memory often exceeds 1 MB.

For example, when a PC sends a read command to a drive, it first checks for requested data in the cache. If the data is in the cache, it can be accessed quickly and sent to the PC. If it isn't in the cache, the drive will find and read the data from the disk and transfer them to the cache. In fact, the drive will read more data than is requested, thus trying to guess the next possible PC request. This guess is typically quite good, since most applications process data sequentially and it is likely that the next request will be made for the data that follow the last request. This type of cache is referred to as *read cache* or *look-ahead cache*.

When the data is written on the disk, a similar approach can be used to speed up system performance. When the PC sends some data to the drive, it then waits for a confirmation that the task is completed. Since the data writing may take some time for a large block of data, this may slow the PC down significantly. Instead, the drive will place those data in the cache (which is done very quickly) and immediately respond to the PC that the task is completed, even if the data isn't stored on the disk yet! Then, a PC can forget about this task and proceed to other tasks. Thus, a large cache is clearly beneficial for the drive and PC performance, and is a good investment.

MICROSOFT CORP. EXHIBIT 1026
Page 11 of 27

Formatting

Low-Level Formatting

Formatting is the first step in making the drive ready for data storage and retrieval. At this stage, the drive is being physically divided into tracks and sectors.

Low-level formatting stays unchanged for the entire life of the drive unless the drive is reformatted. Nowadays, drives are usually sold with low-level formatting already done.

Partitioning

Partitioning divides the drive into logical drives (C:, D:, E:, etc.). Every drive has at least one "primary partition" (C:) and may have many extended partitions. The primary partition contains drive-booting information in the Master Boot Record (MBR) and also keeps a record of all other partitions. A partition is usually made using the FDISK.exe program. The following procedure is typically used to partition the drive on a PC:

1. Remind yourself that all information present on the drive will be destroyed if you **complete step 2**!
2. Type "fdisk" command in a DOS window and press Enter.
3. Select "Create DOS partition or logical DOS drive" by pressing 1. Then press Enter.
4. Select "Create Primary DOS partition" by pressing 1. Create your first drive partition. To make this partition bootable, specify the partition as *active*.
5. Verify that all drive space was partitioned. If needed, create an extended partition with extra logical drives.

If the FDISK program doesn't work with the drive, type "fdisk/status" to check if your drive is present.

High-Level Formatting

High-level formatting prepares drive partitions for the operating system by creating a root directory, from which all other subdirectories can be created, and creating a File Allocation Table (FAT), which keeps track of all information on the disks and all the relationships between different pieces of information. A loss of the FAT is equivalent to loss of data, since the system will not be able to attribute data to the specific files

even if the data themselves are intact. This operation is usually done using the FORMAT.exe program. The following procedure is typically used to format a drive on a PC:

1. Remind yourself that all information present on the drive will be destroyed if you complete step 2 below using the **wrong drive letter**!
2. In the DOS window, type "format x:/s," where x: is the letter of your first partition. Repeat the formatting process for all new partitions.

File System

The file system is a high-level environment allowing the user to interact with the data stored in files on various storage systems. It allows the user to actually address data as files by keeping track of the file location, name, length, and so on.

There are three primary file systems on today's IBM-based PC:

- FAT16 (DOS and Windows 3.x)
- FAT32 (Windows 95 and higher)
- NTFS (Windows NT)

The main problem with FAT16 is that it is unable to address more than 2.1 GB of data on each logical drive of the hard disk—a small number by today's standards. This means that if you use a FAT16 file system with a larger drive, you will be forced to partition it with the FDISK program to logical disks no larger than 2 GB. Otherwise, you will lose some of the drive's capacity. This is getting harder to do with modern disk drives of 20 GB in capacity and more. Another problem is with the way FAT16 divides sectors into clusters: the cluster size is only 2 KB for logical disks less than 128 MB and 32 KB when the disk size exceeds 1 GB. This means that the system will allocate 32 KB for even the smallest files of 1 KB.

There is no problem with addressing data with FAT32 and NTSF systems. At least, problems won't appear until about 4 TB (terabytes) of storage capacity. The minimum cluster size for FAT32 is reduced to 16 KB. The main problem here is that the FAT32 and NTFS systems cannot read each other's disk partitions, which requires reformatting (FDISK and FORMAT) the drive when the system is changed.

Note:

Windows 98 allows for changing an older, FAT16 file system to a newer 32-bit FAT32 using a built-in program Drive Converter (FAT32), which can be found in START → PROGRAMS → ACCESSORIES → SYSTEM TOOLS → Drive Converter (FAT32).

MICROSOFT CORP. EXHIBIT 1026
Page 13 of 27

e Systems

ke HDD)
s to ran-
for the

e a request
e time the
ve;
y, and the

n over a

y is better
ransfer.

orted
Chapter 2)

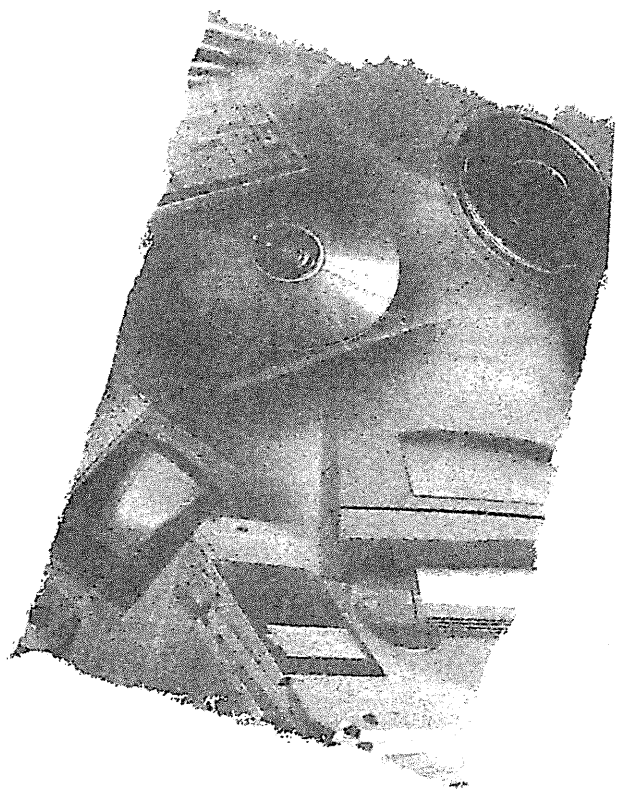
) media

edia

6 Nonvolatile Solid-State Memory

In this chapter...

- Introduction *218*
- Basics of Semiconductor Memory Technology *221*
- Computer Random Access Memory (RAM) *225*
- Flash Memory *229*
- Flash Memory Card Formats *235*
- Solid-State Hard Disk Drives *239*
- Summary *241*



MICROSOFT CORP. EXHIBIT 1026
Page 14 of 27

INTRODUCTION

Memory is the place where a computer can store data and programs and execute them. PC Main memory is a random access solid-state memory (RAM), which enables the computer's CPU (central processing unit) to access instructions and data at a high speed. The reason for the CPU to keep all the important information in RAM is that solid-state memory is much faster than all other computer data storage technologies, such as the hard disk drive, CD-ROM, DVD, and so on.

As was mentioned in Chapter 1, people sometimes confuse the terms *storage* and *memory*, since both of them are used to describe computer devices that store information.

One difference here is that *storage* is designed for less frequently used data or data that could be accessed at a *relatively slow speed*. The *memory* is the place where the computer keeps its most needed and frequently used data, which it can access at a much *faster speed*. An analogy with the file cabinet (storage) and the work desk (memory) is sometimes used to illustrate this difference. One keeps the materials he uses frequently right in front of him on top of the desk, while all other materials are kept in a large file cabinet. Clearly, if one keeps too much data on the desk, it becomes increasingly difficult to locate the needed item, and the data access speed decreases. This is one reason why computer memory is typically much smaller than storage.

Another difference between storage and memory is that storage is *always nonvolatile*, while memory is *mostly volatile*. This means the following: when the power is turned down, nonvolatile memory (hard disk drive, tape, CD, Flash chip, etc.) retains all the information stored while volatile memory (DRAM, SRAM, etc.) loses all the information stored.

In some border cases, a device used for data storage can still be called "memory," thanks to tradition. For example, in the Palm Pilot personal digital assistant (PDA), there are two distinct types of semiconductor memory used: larger volatile RAM and smaller nonvolatile ROM (read only memory). The data in RAM are available for as long as the power is supplied to the system. The data in ROM—which is based on Flash memory technology—is permanently stored and is independent on electric power. The ROM stores the program code, which becomes available upon request and expands into RAM for high availability and fast access. Therefore, the semiconductor ROM here performs a function of a storage device while another semiconductor memory is used purely as the memory device.

Sometimes, the same function can be performed by both the semiconductor memory device and another type of storage device. For example, in some of the newest personal digital assistants (PDAs, see Chapter 7), the same function of ROM can be performed by either the Flash *memory* card or by a small hard disk drive, which is a magnetic *storage* device.

MICROSOFT CORP. EXHIBIT 1026
Page 15 of 27

When thinking about "perfect" computer memory, it is relatively easy to assemble a list of desired characteristics:

- Nonvolatility
- High storage density (high capacity/small volume)
- Fast read, write, and erase operations
- In-system rewritability (by the user, not by the manufacturer only)
- Durability
- Low power consumption
- Low cost
- Single-power supply (for solid state memory)
- Scalability (easy to upgrade to a higher capacity)

In reality, there is always a compromise between the capacity, data access rate, and volatility, as was shown earlier (see Chapter 1) in the computer memory and storage hierarchy. Another important parameter here is always cost.

The solid-state memory is typically very fast, with an access time of 50 to 100 ns compared to more than 5 ms for a hard disk drive, and more than 100 ms for a CD-ROM drive. But the cost one pays for this speed gain is on the order of a magnitude higher (per unit storage capacity) than that of a hard disk or a CD-ROM. This is why different tasks require different types of memory and storage.

There is a large variety of semiconductor memories that differ in speed, cost, storage density, writing/rewriting methods, and volatility. Table 6–1 compares memory and storage attributes for some popular semiconductor memories and magnetic storage devices.

Table 6–1 Comparison of Memory and Storage Attributes

Attribute	Flash	DRAM	SRAM	EEPROM	UV-EPROM	Hard disk drive	Floppy disk
Nonvolatility	Y			Y	Y	Y	Y
High density	Y	Y			Y	Y	
Low power	Y			Y	Y		
In-system rewritability	Y	Y	Y	Y		Y	Y
Fast read	Y	Y	Y	Y	Y	Y	
Fast write	Y	Y	Y	Y	Y	Y	

MICROSOFT CORP. EXHIBIT 1026
Page 16 of 27

Table 6–1 (continued)

Attribute	Flash	DRAM	SRAM	EEPROM	UV-EPROM	Hard disk drive	Floppy disk
High endurance	Y	Y	Y	Y		Y	Y
Low cost	Y	Y			Y	Y	Y
Single-power supply	Y	Y	Y	Y		Y	Y

Figure 6–1 illustrates the features of some solid-state memory technologies.

It seems that Flash memory is best positioned and combines high storage density with nonvolatility and electrical rewritability. This is why this chapter of the book *reviews* computer RAM memory, but *concentrates* on the nonvolatile solid-state (or semiconductor) memory, namely, Flash memory.

From many points of view, Flash memory is the perfect medium for nonvolatile and rewritable storage. Like all other types of solid-state memory, Flash doesn't have moving parts, is compact, is tolerant to a high degree of mechanical shock and vibration, and has low energy consumption.

However, magnetic and optical disk drives still offer much higher cost efficiency, with Flash memory being cost-effective for some specialized applications only, where lost time translates into lost human life or money. In addition, the data rewrite operation within Flash memory is much slower than the read operation, and is slower than that in the modern hard disk drives.

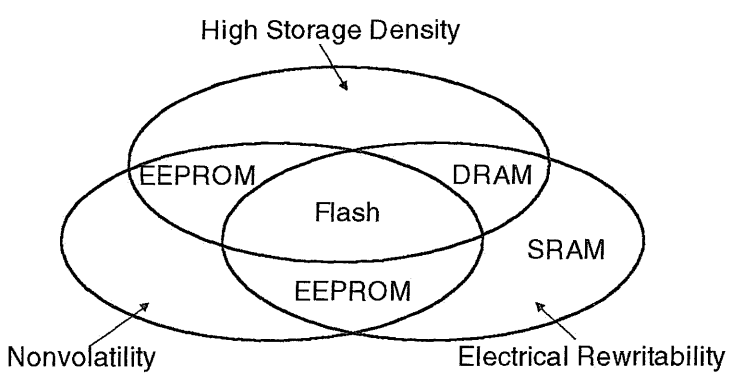


Figure 6–1
Features comparison for some solid state technologies.

MICROSOFT CORP. EXHIBIT 1026
Page 17 of 27

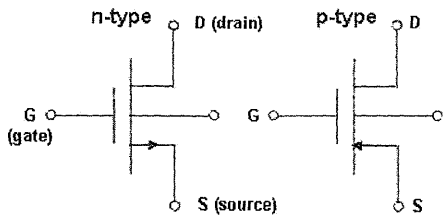


Figure 6–6
The typical schematic representations of the NMOS and PMOS transistors.

COMPUTER RANDOM ACCESS MEMORY (RAM) . . . . .

Let's first look at the part of a computer where the real action takes place: the cache and the Main memory.

In order for the central processing unit (CPU) to run an application, this application must reside in the memory, which is called the Main memory or System memory. When a specific application is needed for an operating system, the system will find it on the storage device (*i.e.*, hard disk drive) and then transfer it to Main memory. Figure 6–7 shows the memory structure in a computer system with high-speed cache memory (L1) integrated into the processor, high-speed secondary cache (L2) included in the system to boost performance, and a relatively large Main memory, which is significantly slower and has a large access time penalty.

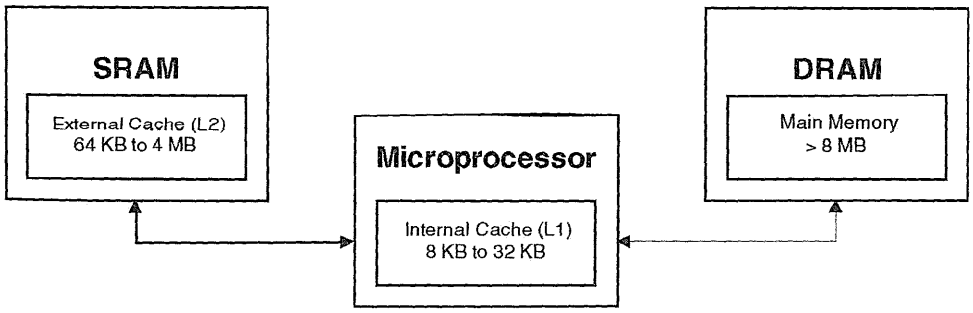


Figure 6–7
A memory hierarchy with high-speed cache memory (L1) integrated into the processor, high-speed secondary cache (L2) included in the system to boost performance, and a relatively large and slower main memory.

MICROSOFT CORP. EXHIBIT 1026
Page 18 of 27

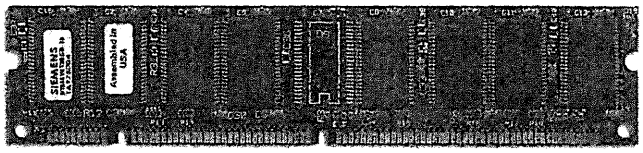


Figure 6–11
168-pin DIMM. Current memory module design. (5.37" x 1.0")

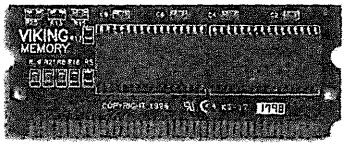


Figure 6–12
72-pin SODIMM. (2.375" x 1.0")

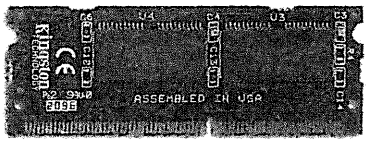


Figure 6–13
144-pin SODIMM. (2.625" x 1.0")

There are two different types of 168-pin DIMM modules: the buffered memory modules and nonbuffered memory modules. The buffered memory modules represent the first generation of 168-pin DIMMs and have memory chips mounted horizontally on the printed circuit board (PCB).

Laptop computers require physically smaller memory modules. Those found in laptop computers are 72-pin and 144-pin SODIMM modules, shown in Figures 6–12 and 6–13.

Even cache memory is now available in modules called Cache-On-A-Stick (COASt).

FLASH MEMORY

The potential of metal-oxide semiconductors (MOS) technology to lead to high density and high performance of the memory devices became known in the mid-1960s. Then, chip makers started thinking hard on how to solve the main problem associated with the MOS memory concept—its volatility.

MICROSOFT CORP. EXHIBIT 1026
Page 19 of 27

The Main memory (RAM) is most responsible for handling operating and application needs. Unfortunately, being a *volatile memory*, RAM loses all data when the system is powered down.

Another type of memory is called ROM, or Read Only Memory. ROM has data permanently stored even when the system is off, and represents a *nonvolatile* semiconductor memory (NVSM) technology.

The first solutions for the volatility problem came in 1967 in the form of a floating gate concept and the metal-nitride-oxide-semiconductor (MNOS) memory device. An ultraviolet (UV)-erasable programmable ROM (PROM/EPROM) of 1 Kb became widely available in 1971. In 1983, 16 KB EEPROM were introduced.

There are three major types of ROM: PROM (Programmable ROM), EPROM (Erasable PROM), and EEPROM (Electrically Erasable PROM).

PROM cannot be changed after being recorded once. EPROM allows the manufacturer to remove one set of instructions in exchange for another set. EEPROM is upgraded by "flashing" the chip and is called Flash ROM.

Flashing is done with the help of a special software program that stores new data on the chip. The term "Flash" refers to the fact that the entire content of the memory chip is erased in one step. In fact, current generation of Flash memory devices differs from both EEPROM and EPROM in that the "erase" operations of the Flash chip can be done in blocks.

Flash, as well as EPROM and EEPROM, must be erased before it is written to. Whole EPROM chips are, for example, erased at once by UV light. EEPROM is erased automatically before the "write" operation. The Flash chip can be erased all at once (bulk erase flash) or in blocks (boot block erased flash).

Therefore, Flash memory is both *nonvolatile* and *easy to reprogram*. This is why Flash memory is often used for data storage in digital cameras and camcorders (still image option), digital cellular phones, PC cards for laptop PCs, audio digital recorders, PDAs, GPS, answering machine sets, and so on (see Chapter 7).

Basics of Flash Memory Technology

All nonvolatile semiconductor memory devices store the data in a form of electric charges in the gate insulator of a MOSFET.

As was mentioned earlier, Flash memory is called so because the entire sections of the microchip are erased at once (or flashed). The erasure is caused by the so-called Fowler-Nordheim tunneling of electrons from the floating gate back to the substrate during the *erase* cycle.

Flash memory devices (see Figure 6–14) is built on a poly-silicon structure floating-gate MOSFET. The control gate is connected to the *wordline* and is isolated from the

MICROSOFT CORP. EXHIBIT 1026
Page 20 of 27

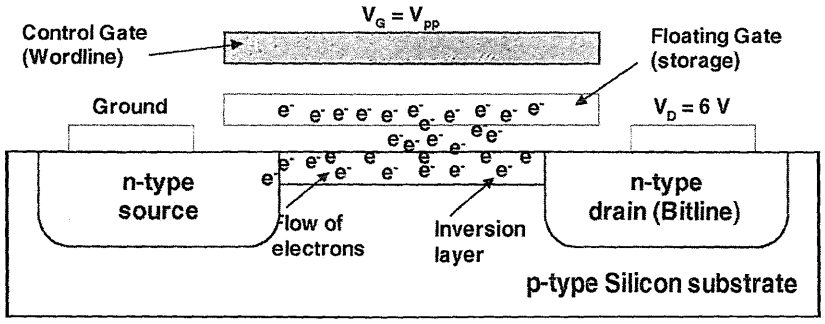


Figure 6–14
An illustration of the Flash cell writing process.

floating gate by the dielectric layer. The floating gate, which is the effective storage medium in a flash device, is isolated from the substrate by a thin oxide layer of about 10 nm thick. This isolation is needed to enable the floating gate layer to store charges.

Flash write operation is based on *hot electron injection* into the floating gate when the *high programming voltage* (V_{pp} = 12V) is applied to the control gate. This voltage creates the inversion region in the p-substrate and allows the electrons from the source to start flowing toward the drain. The electrons are then accelerated by the drain voltage and gain enough energy to "jump" or be "hot-injected" into the floating gate.

After the write cycle is completed, it forms a negative charge on the floating gate, which raises the flash cell's threshold voltage above the so-called wordline *logic 1 voltage*, used by a sense amplifier for detection of the cell current. If the flash cell is written, it corresponds to a "0," and the cell with no negative charge on the floating gate is interpreted as "1." When the wordline of this cell is brought higher than the threshold voltage during read operation, this cell will not "turn on" and conduct the current, thus allowing interpretation of its state as "1."

Flash erase operation (see Figure 6–15) uses the Fowler-Nordheim electron tunneling to remove the charge from the floating gate by applying *high programming voltage* (V_{pp} = 12V) to the source. The high positive voltage on the source makes the electrons move from the floating gate to the source. This changes the state of the cell back to "1." When the wordline of this cell is brought higher than the threshold voltage during read operation, this cell will "turn on" and conduct more current, thus allowing interpretation of its state as "0."

Flash memory cells are organized arrays and are connected via the wordlines and bitlines.

During the *read* operation, the wordline voltage is brought to a voltage higher than the threshold voltage of the erased cell. This turns the erased cells on, which allows for current to flow from drain to source. At the same time, all written cells (with

MICROSOFT CORP. EXHIBIT 1026
Page 21 of 27

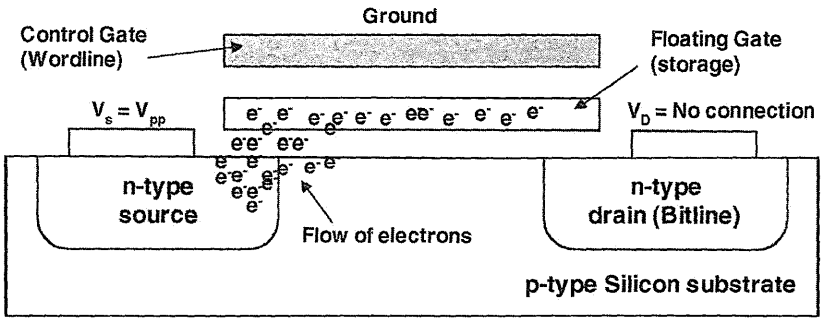


Figure 6–15
An illustration of the Flash cell high-voltage source erase process.

electrons "sitting" on the floating gate) will stay in the off state with almost no current flowing from drain to source. The cell current is then detected by the sense amplifier, allowing interpretation of its state.

Flash Memory Cards

Flash memory card loses power when it is disconnected from the PC, yet the data stored in it is retained indefinitely, until it is rewritten. Flash memory cards use flash memory chips to store and retrieve data. There are two basic types of cards: PC cards (linear cards) and ATA Flash PC cards.

PC cards, also called linear cards: PC cards are typically connected directly to the PC bus and have direct access to the CPU (or CPU has direct access to the card). This mode of operation is called "execute in place" (XIP) mode and is a method of directly executing the applications from the card's ROM without loading the application into DRAM. XIP requires high-speed random access during read operation. This type of application is typically needed for internal PC storage of the system data, the boot code, BIOS, fonts, and so on.

ATA Flash PC cards: PC cards that are connected to the CPU via the ATA interface, similar to that used by other storage devices such as hard disk drives. The interface could either belong to the computer or to the PC card. CPU in this case treats the ATA Flash PC card just as a solid-state equivalent of the hard disk drive (HDD) and the card must be capable of addressing (rewriting) data using 512-byte blocks—the way HDD does it.

Figure 6–16 shows how the PC cards of two different types communicate with the CPU. Linear cards communicate directly with the CPU in XIP mode. ATA Flash PC cards use DRAM (Dynamic Random Access Memory) to communicate with the CPU.

Generally, the linear cards use DINOR and NOR flash chips logic, while ATA PC cards use AND and NAND flash chips logic. Those chips are electrically pro-

MICROSOFT CORP. EXHIBIT 1026
Page 22 of 27

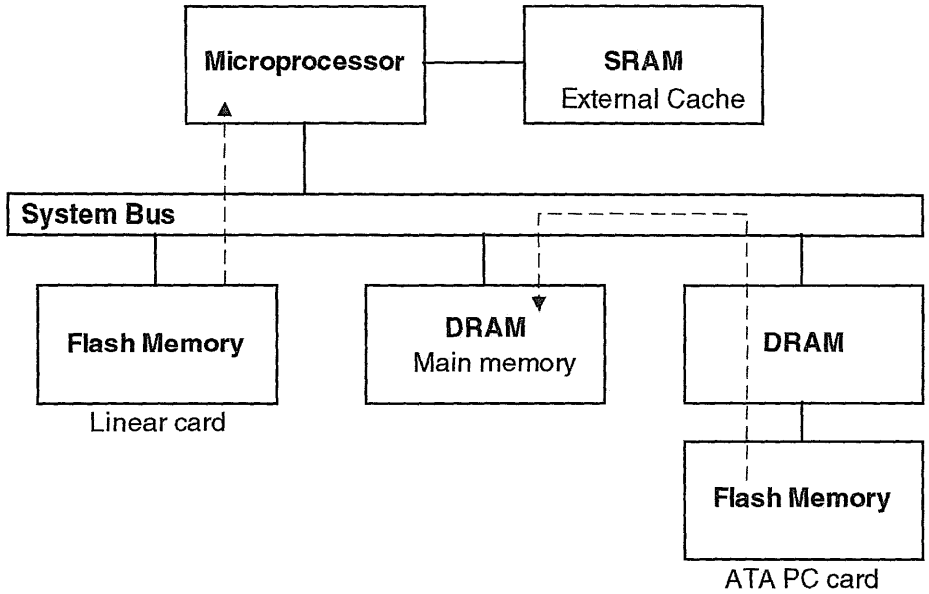


Figure 6–16
PC Card communication with the CPU.

grammable and reprogrammable nonvolatile solid-state data storage devices. Next is the explanation to the different logical designs used in Flash cards.

NOR

The logic of NOR is shown next and should be understood as follows (for the first row):

```
If A = 0 and B = 0, then C = 1.
```

A	B	C
0	0	1
0	1	0
1	0	0
1	1	0

NOR represents the earliest flash design. NOR flash memory has a large sector size and cell size, high-speed rewrite, and high-speed random data access. NOR devices use energy consuming hot electron injection for data writing and tunnel release for erasing. One disadvantage of NOR is that they are difficult to scale up to higher capacities.

MICROSOFT CORP. EXHIBIT 1026
Page 23 of 27

NAND

The logic of NAND is shown next:

A	B	C
0	0	1
0	1	1
1	0	1
1	1	0

NAND flash chips are serial access devices with high storage capacity, single voltage (3.3V or 5V) operation, small block size, small cell size, and high durability. NAND use tunnel injection for writing and tunnel release for erasing, which is a power-saving technique.

DINOR (Divided Bit-Line NOR)

DINOR flash chips (introduced by Mitsubishi) have high random data access, single voltage supply (3.3V or 2.7V), fast block erase, large block size, multi-sector erase capabilities, and low power consumption using tunnel injection/tunnel release for write and erase. Storage capacity (called density) for DINOR is currently in the range from 1 MB to 2 MB.

AND

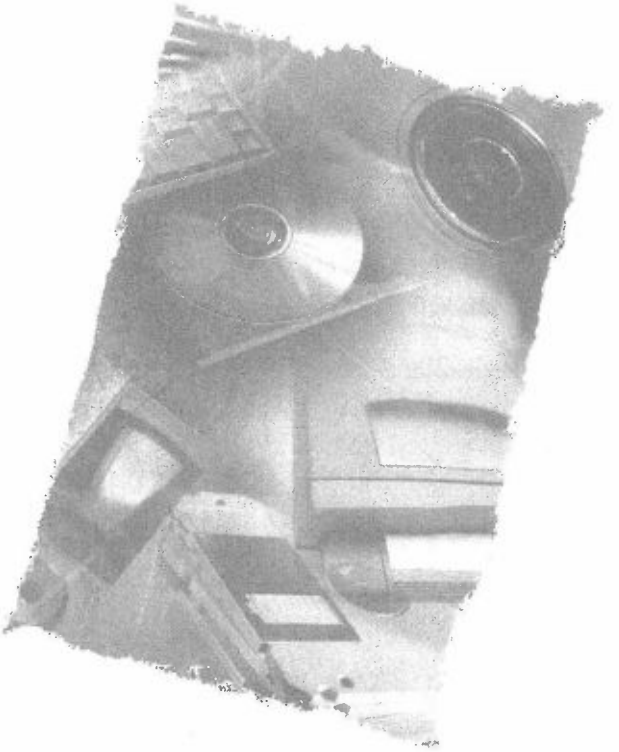
The logic of AND is shown next:

A	B	C
0	0	0
0	1	0
1	0	0
1	1	1

The AND Flash chip (first introduced by Hitachi) is a serial access device combining the best of NOR and NAND devices. It has fast erase, single voltage supply operation, very high capacity, low power consumption (writing by tunnel injection and erasing by tunnel current release), and small block size. It offers multiblock erase and auxiliary byte redundancy for ECC. AND flash chips are optimized for high-capacity storage devices such as ATA Flash PC cards. Devices with more than 32 MB of storage capacity are available.

MICROSOFT CORP. EXHIBIT 1026
Page 24 of 27

Glossary



Actuator
The internal part of the drive that moves the read/write element from one location to another. Today's hard disk drives are equipped with voice-coil actuators (VCA) that are much faster than the previously used step motor. A typical drive's VCA consists of a rotary voice coil and the element-mounting arm. As current is applied to the rotor, it rotates at frequencies of 50 Hz or more, positioning the heads over the desired cylinder on the media.

Access time
See average access time.

Adaptive Battery Life Extender
A technology used in mobile hard disk drives to reduce power consumption by approximately 20 percent. It is accomplished by adapting power mode transitions based on the access patterns of individual users, rather than relying on conventional timers. As a result the power consumption may be lowered and battery life increased.

ADC
Analog-to-digital conversion. A multistage process of converting analog signal into its digital equivalent.

ADSL
Asymmetric DSL. ADSL is one of the most popular DSL technologies of today. Usually, the "A" is dropped, so, when we talk about DSL for home use, we're talking about ADSL.

AFM
Atomic Force Microscope. Physical tool for measurements of nanoscale topography and forces. Basically, the AFM consists of a thin, short lever with a sharp tip at the end. The tip's radius is usually made on the order of a few nanometers. When this tip is brought in proximity to the surface, surface forces act on the tip and result in bending and torsion of the lever, which is measured (mostly optically).The bending and torsion provides for a measure of the topography and forces.

Air-bearing slider
A small ceramic slider with a self-pressurized air bearing and magnetic read-write element at the end. When a magnetic rigid disk starts spinning, the lift force is generated between

MICROSOFT CORP. EXHIBIT 1026
Page 25 of 27

the slider and disk, which makes the slider fly at a very small height of a few dozen nanometers. This slider design was first used in the IBM "Winchester"-type hard disk drives that revolutionized drive technology.

AIT
Advanced Intelligent Tape. Relatively new 8-mm magnetic tape format.

AME
Advanced Metal Evaporated (ME) tape. The newer generation of ME tapes.

AMP
Advanced Metal Particle (MP) tape. The newer generation of MP tapes.

ANSI
American National Standards Institute represents the United States in the International Organization for Standardization (ISO) and works on coding and signaling standards.

Arbitrated loop
When devices form a Fibre Channel Arbitrated Loop, the transmitters are connected to a receiver for all nodes involved. When a node port "wins" arbitration, it establishes a point-to-point connection with another node.

Areal density
Is an important indicator of the disk (and drive) performance. Corresponds to the number of magnetic bits per unit area, and is usually measured in Gb/in^2.

Asynchronous Data Transfer
The data transfer scheme that allows packets of data to be transferred only after receiving a response from the receiving device. The retries are mandated if errors occur. This method sacrifices transfer speed for reliability. *See also* synchronous data transfer and isochronous data transfer.

ATAPI
Advanced Technology Attachment Packet Interface. A type of interface used to connect additional hardware devices to a computer. This interface is commonly used to connect internal devices such as CD-ROMs, hard drives, tape drives, and Zip drives. The advantages of ATAPI are that it is economical and offers high performance. To install an ATAPI device you must open your computer case and be able to identify the IDE channel(s) on your computer's motherboard.

ATA/ATA-1
Advanced Technology Attachment. A type of interface used to connect additional internal hardware devices to a computer. ATA is the official name given by ANSI to the IDE interface.

ATA-2/Fast ATA-2
ATA-2 is also known as enhanced IDE. Fast ATA-2 includes upgrades in DMA and PIO modes and offers higher DTR and some new commands.

ATA-3
An update to the ATA-2 standard and includes SMART capability. ATA-3 doesn't offer higher speed but adds new commands and more precisely defined procedures.

ATA-4
ATA-4 is also known as Ultra ATA, UDMA, or ATA-33. An update to the ATA-3 standard, it includes Ultra DMA (direct memory access) capability. Supports the maximum DTR of 33 MB/sec (in burst mode) and provides for enhanced reliability and data integrity through the use of double-edge clocking and Cyclical Redundancy Checking (CRC).

ATA-5 (or ATA-66)
One of the newest and fastest IDE interfaces that doubles the maximum DTR of ATA-33

MICROSOFT CORP. EXHIBIT 1026
Page 26 of 27

to 66 MB/sec (in burst mode). Provides for enhanced reliability and integrity through CRC and the use of newer 40-pin 80-conductor cable instead of the 40-pin 40-conductor cable. The ATA-5 drives are completely backward-compatible with the previous ATA drives and host systems.

ATA-6 (or ATA-100)

This is the newest and the fastest IDE interface at the moment. Supports the maximum DTR of 100 MB/sec (in burst mode). Provides for enhanced reliability and integrity through CRC. Uses a new 40-pin 80-conductor cable instead of the 40-pin 40-connector cable. IDE connectors and the ATA-6 drives are completely backward-compatible with the previous ATA drives and host systems.

ATOMM

Advanced super Thin layer and high-Output Metal Media. The latest advancement in the world of MP media.

ATM

Asynchronous Transfer Mode. A high-speed switching technique used to transmit voice, data, and video.

Average access time

The average delay within the drive between receiving a write or read command and the instance when the head begins reading or writing. The access time includes seek time, settling time, and latency.

Backup (of data)

Important procedure of saving data on a separate data storage device to prevent complete data loss in case of unexpected failure of main storage system. A daily backup is highly recommended for active computer users.

Bad block

A block of data (usually the size of a sector) that cannot reliably hold data because of a media flaw or damaged format markings.

Bandwidth

The amount of data that can be transmitted in a given amount of time. Bandwidth is also the difference between the highest and lowest frequencies of the signal used. Usually expressed in bits per second (bps) or in Hz (MHz, GHz).

Benchmark

A specialized set of programs used to evaluate the performance of a computer or one of its elements, such as a drive or memory.

Bezel

A decorative cover for the drive and is also known as the faceplate.

Block

A group of bytes stored and accessed as a logical data unit. One block of data is typically stored on the hard disk as one physical sector of data.

BIOS

Basic Input/Output System. The BIOS is built-in software that defines what a PC does without accessing programs from a disk. It makes it possible for a computer to boot itself. On PCs, the BIOS contains all the code required to control the disk drives, keyboard, display, serial communications, and some other functions. The BIOS is typically found on a ROM chip that comes with the computer. Since RAM is faster than ROM, many computer manufacturers make systems that copy BIOS from ROM to RAM every time the computer is restarted. This process is called shadowing.

BIOS Flash/Upgrade

BIOS Flash or BIOS Upgrade is the replacement of a current BIOS with a newer version having increased capabilities. Usually done

MICROSOFT CORP. EXHIBIT 1026
Page 27 of 27