**Chapter 6: External Memory**

Table of Contents

[6.1 Magnetic Disk 3](#_Toc48983857)

[Magnetic Read and Write Mechanisms 3](#_Toc48983858)

[Data Organization and Formatting 4](#_Toc48983859)

[Physical Characteristics 8](#_Toc48983860)

[Disk Performance Parameters 10](#_Toc48983861)

[6.2 RAID 13](#_Toc48983862)

[RAID Level 0 14](#_Toc48983863)

[RAID Level 1 15](#_Toc48983864)

[RAID Level 2 16](#_Toc48983865)

[RAID Level 3 17](#_Toc48983866)

[RAID Level 4 18](#_Toc48983867)

[RAID Level 5 19](#_Toc48983868)

[RAID Level 6 20](#_Toc48983869)

[6.3 Solid State Drives 21](#_Toc48983870)

[SSD Compared to HDD 21](#_Toc48983871)

[SSD Organization 22](#_Toc48983872)

[Practical Issues 23](#_Toc48983873)

[6.4 Optical Memory 25](#_Toc48983874)

[Compact Disks 26](#_Toc48983875)

[CD and CD-ROM 26](#_Toc48983876)

[CD-R 28](#_Toc48983877)

[CD-RW 28](#_Toc48983878)

[Digital Versatile Disk 29](#_Toc48983879)

[DVD 29](#_Toc48983880)

[Blu-Ray Disks 30](#_Toc48983881)

[6.5 Magnetic Tape 31](#_Toc48983882)

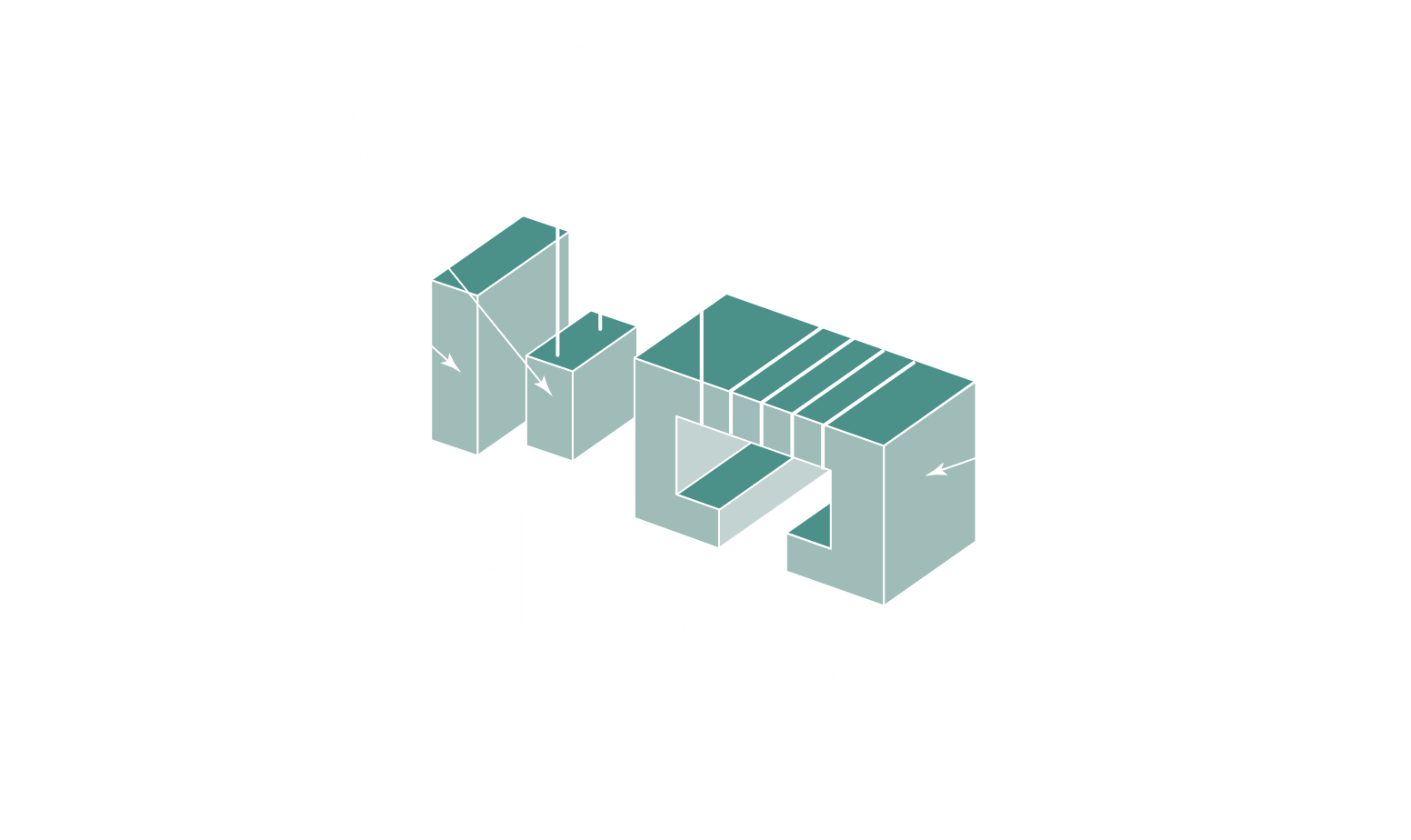
## 6.1 Magnetic Disk

A disk is a circular platter constructed of non-magnetic material, the substrate, coated with a magnetizable material. Traditionally, the substrate was made of aluminium, or an aluminium alloy, but more recently, glass is being used. This is because glass has a number of benefits:

* Improved uniformity of the magnetic film surface, increasing disk reliability
* Significantly reduced overall surface defects, which reduces read-write errors
* Support for lower fly heights (discussed later)
* Better stiffness, to reduce disk dynamics
* Can withstand more shock and damage

### Magnetic Read and Write Mechanisms

Data is written onto and retrieved from the disk using a conducting coil, called the head. Many systems have separate read and write heads. During an operation, the head lies still on the platter, while the disk rotates beneath it.



Electricity flowing through a coil produces a magnetic field. This property is used to write data onto the disk. Pulses are sent to the write head, and the resulting magnetic pulses are recorded on the surface below.

A magnetic field moving relative to a coil causes a current to flow through the coil. This property is used to read data from the disk. The surface of the disk rotates beneath the read head, generating a current in the coil.

The way the read and write mechanisms work, the structure of the read and write heads are the same. This means that a single coil can be used for both operations, instead of two separate ones. Single heads are used in floppy disks and older rigid disk systems.

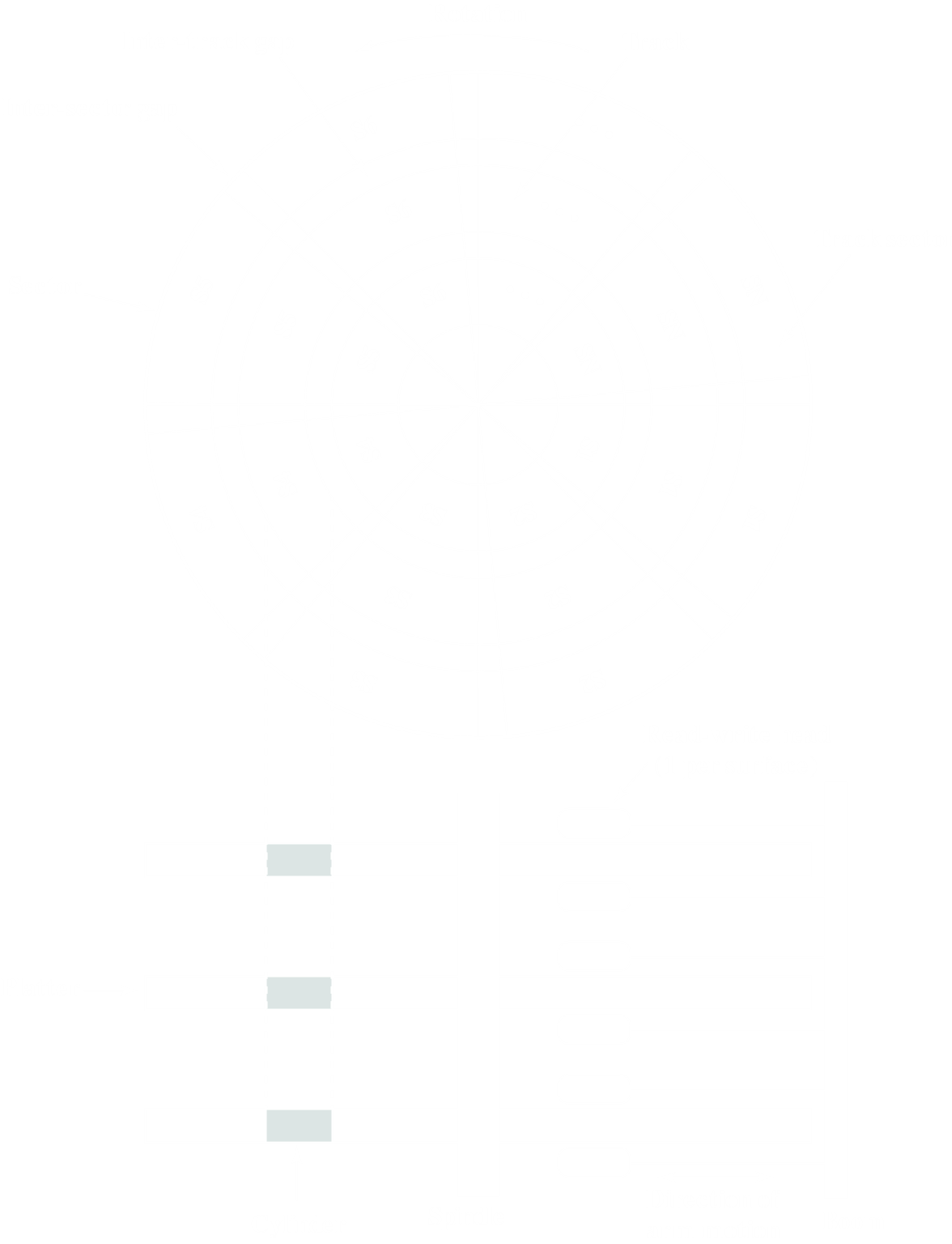
Traditional rigid disk systems use a different read mechanism that requires the read head to be placed close to the write head. The read head has a partially shielded magneto-resistive (MR) sensor, that has an electrical resistance that depends on the direction of the magnetization of the material moving beneath it. By passing a current through the MR sensor, resistance changes are detected as voltage signals. The MR design allows higher frequency operation, which allows greater storage densities and operating speeds.

### Data Organization and Formatting

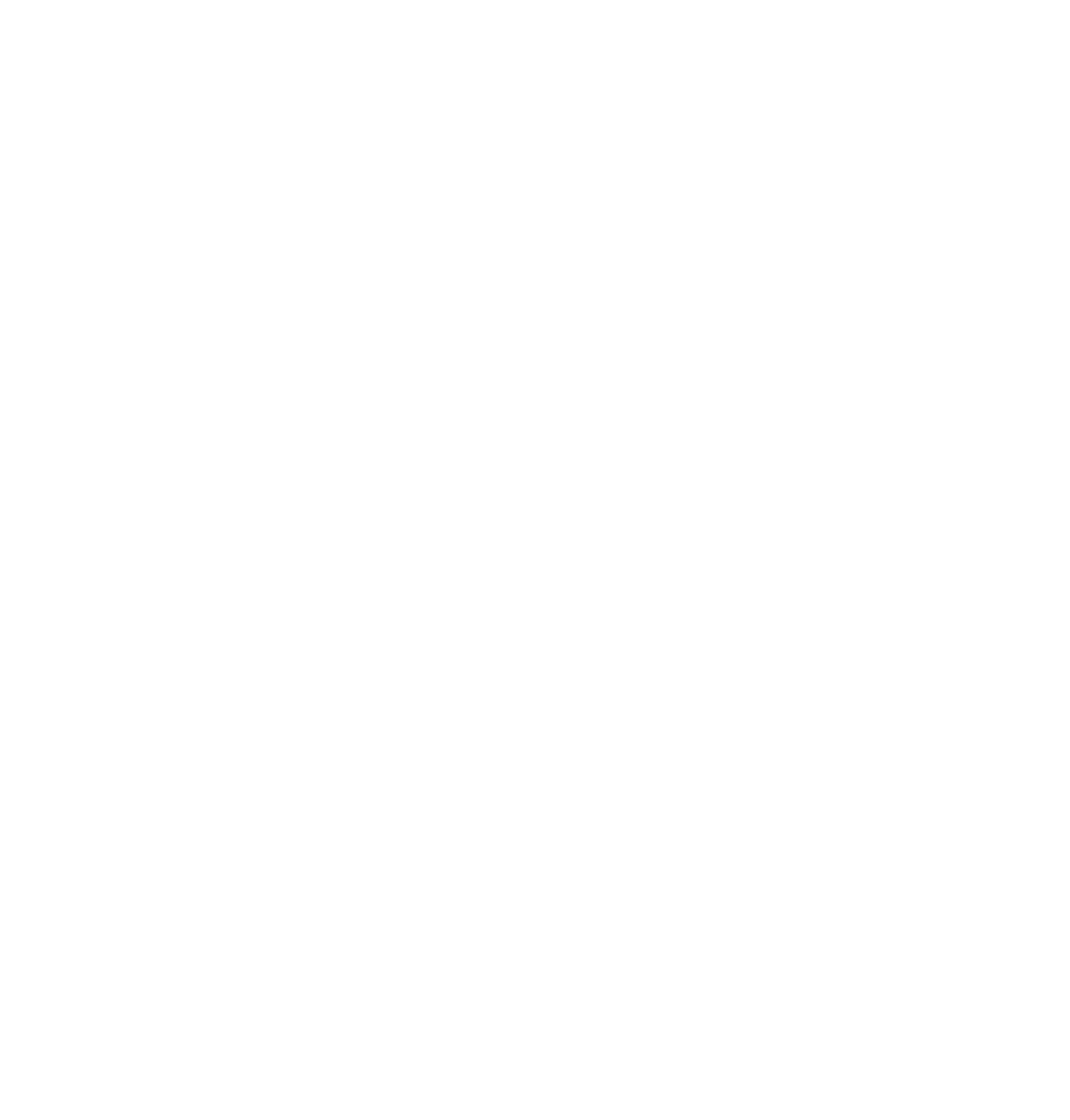
The way the read and write operations work means that data is arranged on the platter in a concentric set of rings, called tracks. Each track is the same width as the head. Adjacent tracks are separated by intertrack gaps. This prevents errors due to misalignment of the head or interference of magnetic fields.

Data is transferred to and from the disk in sectors. The sectors may be of fixed or variable length. Most contemporary systems have fixed-length sectors of 512 bytes. To avoid unreasonable precision requirements, adjacent sectors have intersector gaps.

There are also usually multiple platters stacked together, with data being stored on both sides of each platter, and corresponding heads for each of the sides. Every sector that is on the same line is said to be part of the same cylinder, as shown below. The heads are attached together and rotated together using the boom.



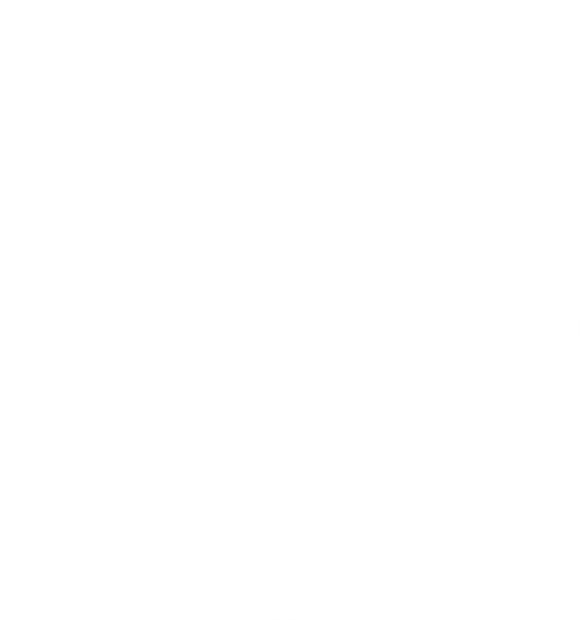
This configuration has a problem. Bits near the centre of the disk travel past the read head faster than bits near the edges. We want to be able to read all the data at the same rate though. In order to achieve this, there is variable spacing between bits of information recorded in different locations, with the outermost tracks having sectors with bigger spacing. Information can then be read at a fixed rated by rotating the disk at a fixed speed, called the constant angular velocity (CAV).



This configuration allows individual blocks of data to be addressed by track and sector. To move the head to a specific address, it must simply be moved onto the right track, and then the disk spun so that the correct sector lies beneath the head.

The disadvantage however, is that the longer outer tracks can only store the same amount of data as the shorter inner tracks. Since density, in bits per linear inch, increases moving from the outermost to the innermost track, disk storage capacity is limited by the maximum recording density of the innermost track. To maximize storage capacity, we need to have the same linear bit density on each track. However, doing so would require extremely complex circuitry.

Modern hard disks use a simple technique that gives approximately equal bit density per track, called multiple zone recording (MZR). The surface of the disk is divided into a number of concentric zones, 16 typically. Each zone contains a number of contiguous tracks, typically in the thousands. Within a zone, the number of bits per track is constant. Zones further from the centre contain more sectors, and thus more bits, than zones that are closer. Zones are defined in a way so that the linear bit density is approximately the same on all tracks of the disk.

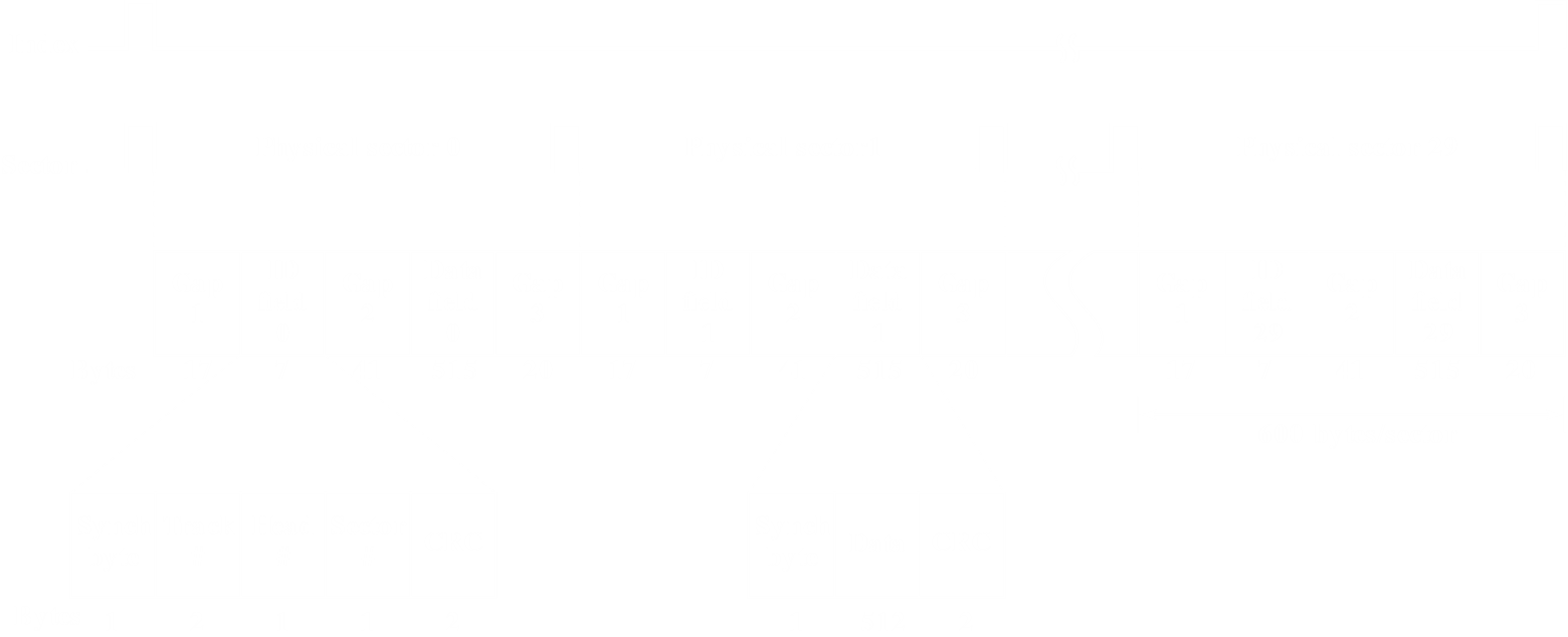


MZR allows greater overall storage capacity in exchange for somewhat more complex circuitry. As the head moves from one zone to another, the length along the track of individual bits changes, causing a change in the timing for read and write operations.

The above figure is a simplified version of MZR. 15 tracks are organized into 5 zones. The innermost two zones have two tracks each, with each track having nine sectors. The next zone has three tracks with 12 sectors each, while the outermost 2 zones have 4 tracks each, with each track containing 16 sectors.

We need to be able to locate sectors, since they are divided in such a way. The start and end of each sector must be identified. This is done with control data recorded onto the disk. Thus, the disk is formatted with some extra data used only by the disk drive and not accessible by the user.

Consider the figure below. Each track contains 30 sectors of 600 bytes each. There are 512 data bytes and control information useful to the disk controller. The ID field is a unique identifier or address used to locate a particular sector. The SYNCH byte is a special bit pattern that delimits the beginning of the field. The track number identifies a track on a surface. The head number identifies a head, because the disk has multiple surfaces. The ID and data fields also contain error-detecting code.



Sometimes, disk sectors get permanently damaged. These are called bad sectors. When a bad sector is found, it is marked so that the operating system skips it in the future. All the data in that sector is lost and read and write operations are no longer possible. Some dedicated software called disk drivers or the operating system itself marks these sectors as occupied or redirects them to another pre-occupied good sector.

### Physical Characteristics

We can differentiate magnetic disks based on several physical characteristics, listed below:

|  |  |
| --- | --- |
| **Head Motion** | Fixed head (one per track) |
| Movable Head (one per surface) |
| **Disk Portability** | Nonremovable disk |
| Removable disk |
| **Sides** | Single sided |
| Double Sided |
| **Platters** | Single platter |
| Multiple platter |
| **Head Mechanism** | Contact (floppy) |
| Fixed gap |
| Aerodynamic |

A fixed head disk has one read-write head on every track, with the arm on which the heads reside extending across all the tracks. They are rare nowadays. A movable head on the other hand, is a single head mounted on an arm that can extend or contract. This allows it to be positioned above any of the tracks.

The disk itself is mounted in a disk drive, which contains the arm, a spindle to rotate the disk and electronics for input and output. A non-removable disk is permanently mounted in a disk drive, such as the hard disk in a personal computer. A removable disk on the other hand, can be replaced with another disk. This allows unlimited amounts of data with a limited number of disk systems. We can also move such disks from one computer to another. Floppy disks and ZIP cartridge disks are removable disks.

Most disks have a magnetizable coating applied to both sides of the platter, and are called double sided. Some cheaper disk systems use single sided disks, which only have the magnetizable coating on one side of the platter.

Some disk drives contain multiple platters stacked vertically, a fraction of an inch apart, which make use of multiple arms. They also use one movable read-write head per platter surface. All the heads are fixed so that they are the same distance from the centre of the disk and move together. Thus, the heads are always positioned over tracks that are of equal distance from the centre of the disk, i.e. they lie on the set of tracks that are in the same relative position on the platter, a.k.a. a cylinder.

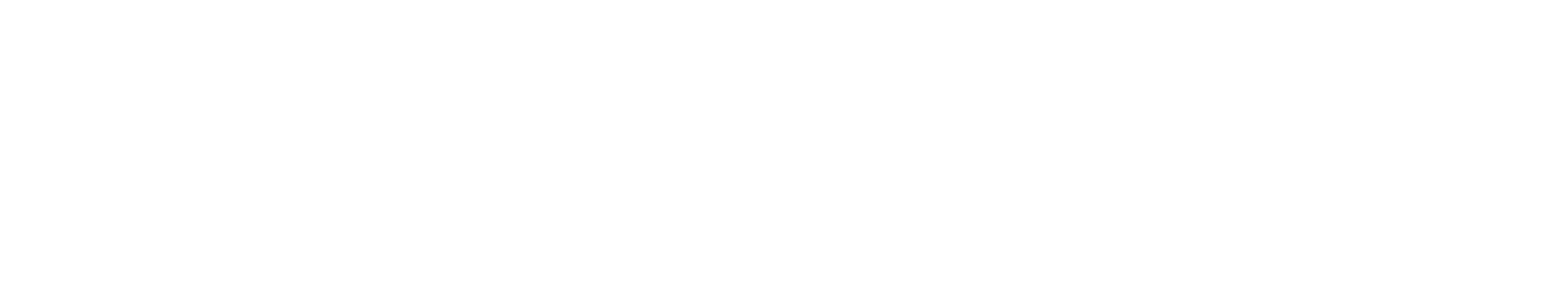
We can also divide disks based on the way the head mechanism interacts with the disk. To write or read properly, the head must generate or read an electromagnetic field of sufficient magnitude. The narrower the head, the closer it needs to be to the platter. A narrower head means a narrower track, which means greater data density, which is good. However, the closer the head is to the platter, the greater the risk of errors due to impurities and imperfections, which is bad. Some disk systems allow the head to come into physical contact with the disk during read or write operations, while others leave a fixed air gap between the head and the disk. A third type, called the aerodynamic gap, is a compromise between the previous two. The Winchester disk was of this third category. They are used in sealed drive assemblies, that are nearly free of contaminants. The head is actually an aerodynamic foil that rests lightly on the platter’s surface when the disk is motionless, and is lifted up by the pressure generated when the disk spins. This resulting non-contact system can be created with narrow heads that operate close to the platter’s surface, allowing greater data density.

### Disk Performance Parameters

The actual details of disk I/O operations depend on the computer system, the operating system and the nature of the I/O channel and disk controller hardware.

When the disk drive is operating, the disk rotates at a constant speed. To read or write, the head must be positioned at the correct track and at the beginning of the correct sector on that track. On a movable head system, the time taken to move the head to the correct track is known as the seek time. This time delay is not present in a non-movable head system. In either system, the time taken to line up the beginning of the sector with the head is known as the rotational delay or rotational latency. The sum of the seek time and rotational delay is the access time, the time taken to get into position to read or write. Once the head is in position, the read or write operation takes place as the sector moves beneath the head, transferring data. The time required for the transfer is known as the transfer time.

There are some other queueing delays associated with disk I/O operations. When a process issues an I/O request, it must wait in queue for the device to be available. If the device shares a single or a set of I/O channels with other disk drives, there may be additional delays after the device becomes available to the process for the required channel to become available.



In some high-end systems for servers, rotational positional sensing (RPS) is used. When the seek command is issued, the channel is released to handle other I/O operations. When the seek is complete, the device determines when the data will rotate under the head. As the sector approaches the head, the device tries to reconnect with the host. If the control unit or the channel is busy with another I/O operation, the reconnection fails. The device must rotate one whole revolution before it can try to reconnect again. This is called an RPS miss. This is an extra delay.

The transfer time is given by,

where is the number of bytes to be transferred, is the number of bytes on a track and is the rotation speed in revolutions per second.

Rotational delay is given by since on average, the delay is half the time taken for one complete revolution.

The total average read or write time is given by

where is the average seek time.

Note that the calculation becomes complicated for zoned drives, where the number of bytes per track varies.

## 6.2 RAID

The rate of improvement in secondary storage performance has been much slower than that of processors and main memory. This mismatch has led to the disk storage system becoming the main focus in improving overall computer system performance.

If a component can only be pushed so far, additional gains in performance may be achieved by using multiple parallel components. In the case of disk storage, this led to the development of arrays of disks that operate independently and in parallel. With multiple disks, separate I/O requests can be handled in parallel, as long as the data required resides in separate disks. A single I/O request can also be handled in parallel if the block of data to be accessed is distributed across multiple disks.

With multiple disks, there are many ways to organize data and add redundancy to improve reliability. This could make it difficult to develop database schemes that are usable on a number of platforms and operating systems. Fortunately, the industry has agreed on a standardized scheme for multiple-disk database design called the Redundant Array of Independent Disks (RAID). The RAID scheme consists of 7 levels, 0 to 6. The levels do not indicate a hierarchical relationship, but rather different design architectures that share three common characteristics:

* RAID is a set of physical disk drives viewed by the operating system as a single logical drive
* Data is distributed across the physical drives of an array in a scheme known as striping
* Redundant disk capacity is used to store parity information, which guarantees data recoverability in case of a disk failure

The details of the second and third characteristics differ for the different RAID levels, with RAID 0 and RAID 1 not supporting the third characteristic at all.

The RAID strategy employs multiple disk drives and distributes data in such a way as to enable simultaneous access to data from multiple drives, thus improving I/O performance and allowing easier incremental increases in capacity. Originally, the term RAID stood for Redundant Array of Inexpensive Disks, in contrast to Single Large Expensive Disk (SLED). The meaning of RAID was later changed when SLED became obsolete.

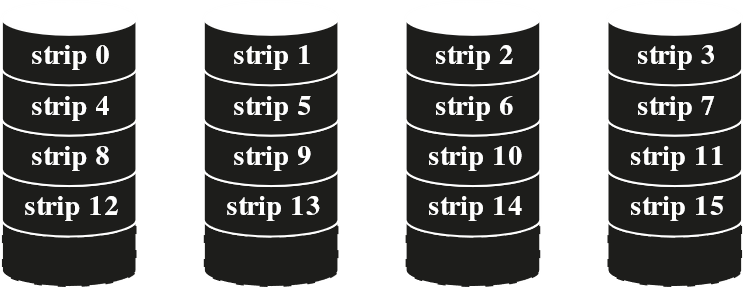
The unique contribution of the RAID proposal is to address effectively the need for redundancy. Although allowing multiple heads and actuators to operate simultaneously achieves higher I/O and transfer rates, the probability for failure increases. To compensate for the decreased reliability, RAID makes use of stored parity information that enables the recovery of data due to disk failure.

### RAID Level 0

This is not truly a member of the RAID family because it does not include redundancy to improve performance. However, there are a few applications, like some on supercomputers, for which performance and capacity are primary concerns and low cost is more important than improved reliability.

For RAID 0, the user and system data are distributed across all of the disks in the array. When two different I/O requests are pending for two different blocks of data, there is a good chance that the requested blocks are on different disks. Thus, the two requests can be issued parallelly, reducing the I/O queueing time.

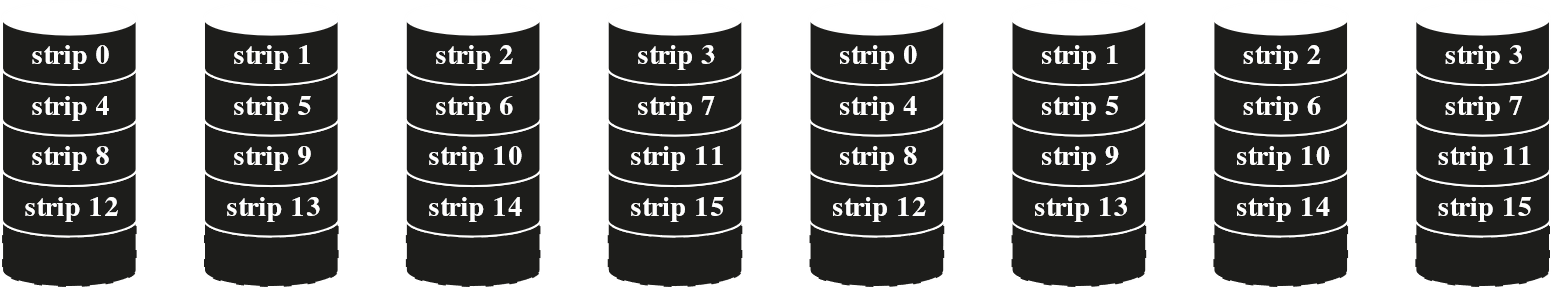
As with all other RAID levels, RAID 0 has something called striping. Consider the diagram below:



All the user and system data are viewed as being stored on a logical disk. The logical disk is divided into strips. The strips may be physical blocks, sectors or some other unit. The strips are mapped as shown to consecutive physical disks in the RAID array. A set of logically consecutive strips that maps exactly one strip to each array member, i.e. the first row in the diagram, is called a stripe. In an n-disk array, the first n logical strips are physically stored as the first strip on each of the disks and form the first stripe, the next n logical strips are physically stored as the second strip on each of the disks and form the second stripe and so on. If the I/O request consists of multiple logically contiguous strips, up to n strips for that request can be handled parallelly, greatly reducing the I/O transfer time.

### RAID Level 1

In RAID 1, the exact same form of data striping is used as in RAID 0, but a second copy of the entire thing is kept to achieve redundancy and improve reliability.



RAID 1 has a few advantages. Read requests can be serviced by either of the two disks that contain the data, whichever has a lower seek time and rotational latency. Write requests require two strips to be updated, but this can be done parallelly. Also, there is no ‘write penalty’, a concept that will become clearer further down. Essentially, since redundancy is not achieved through the use of parity bits, no extra computation needs to be done. Recovery from a failure is simple, since if one drive fails, the data is still available in another drive.

Disadvantages of RAID 1 include a higher cost, since it requires twice the physical disk space of the logical disk that it supports. Because of this, RAID 1 is limited to drives that store system software and data and other highly critical files. In these cases, RAID 1 provides real time copy of all data, so that if a disk fails, all the critical data is immediately available.

### RAID Level 2

RAID 2 and RAID 3 use a parallel access technique. In a parallel access array, all member disks participate in the execution of every I/O request. Typically, the spindles of the individual drives are synchronized so that each disk head is in the same position on each disk at any given time. Data striping is used here as well, but for RAID 2 and RAID 3, the strips are very small, often a single byte or word. With RAID 2, and error-correcting code, typically SEC-DED, is calculated across corresponding bits on each data disk, and the bits of the code are stored in the corresponding bit positions on multiple parity disks.

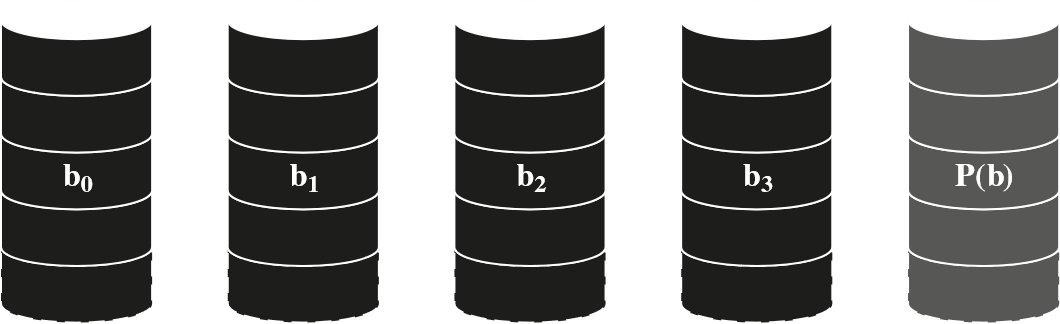


RAID 2 needs fewer disks than RAID 1, but it is still costly. The number of redundant disks is proportional to the log of the number of data disks. On a single read, all the disks are simultaneously accessed and the requested data and error-correcting code are delivered to the array controller. If there is a single error, the controller corrects the error instantly so read access time is not slowed. On a single write, all data disks and parity disks must be accessed.

RAID 2 is only a good choice in an environment where many disk errors occur. Given the high reliability of individual disks and disk drivers, RAID 2 is too good, and is not implemented.

### RAID Level 3

RAID 3 is similar to RAID 2, except that it only has a single redundant disk, no matter how large the disk array. It employs parallel access, with data distributed in small strips. Instead of an error-correcting code, a simple parity bit is computed for a set of individual bits in the same position on all of the data disks. In the event of a drive failure, the parity drive is accessed, and data is reconstructed from the remaining devices. Once the failed drive is replaced, the missing data can be restored on the new drive.



For an array of five drives in which to contain data and is the parity disk, the parity of the th bit is calculated as:

Say the drive fails. Data can be regenerated as:

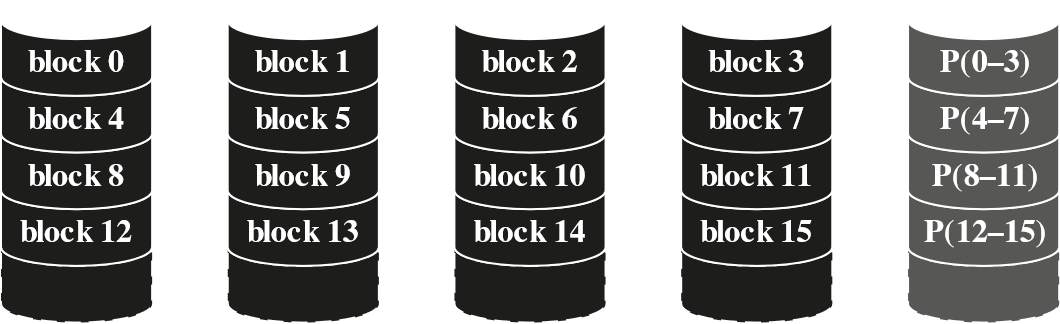
The same method will be used for RAID 4 to RAID 6.

In the event of a disk failure, all of the data is still available in what is known as reduced mode. For reading data, the missing data is regenerated using the calculation above. When data is written, consistency of the parity must be maintained so that the data can later be regenerated.

### RAID Level 4

RAID 4 through 6 use an independent access technique. Each member disk operates independently, so that separated I/O requests can be satisfied parallelly. This makes them suitable for applications that required a large number of I/O requests, and less suitable for applications that require a high data transfer rate.

Data striping is still used, with the strips being relatively large. With RAID 4, a bit-by-bit parity strip is calculated across corresponding strips on each data disk, and the parity bits are stored in the corresponding strip on the parity disk.



RAID 4 has a write penalty when an I/O write request of small size is performed. Each time a write occurs, the parity bits must be updated along with the user data. Initially, say for each bit , this relationship was true:

Say a change is made to disk . This is indicated as . Thus, the new parity bit is:

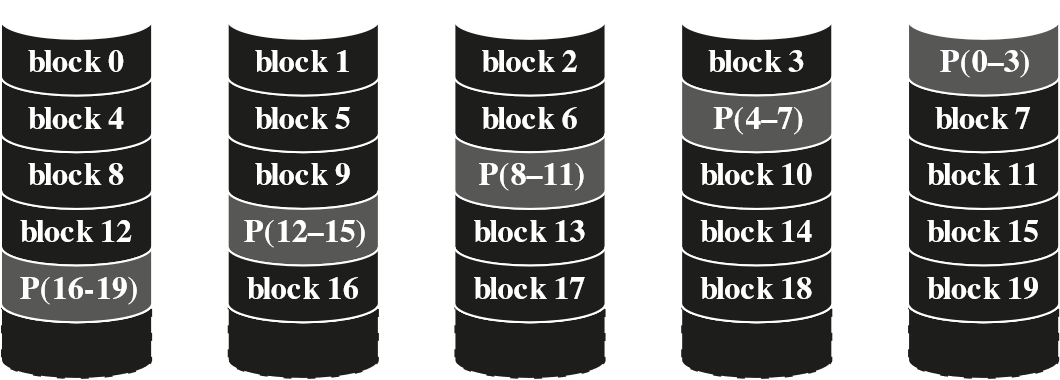
To calculate new parity, the software must read the old user data strip and old parity strip, and update the new user data strip and new parity strip. Thus, each write operation contains two reads and two writes.

In case of large I/O writes that involve strips on all disk drives, parity is simply calculated from the new data bits. There are no extra reads or writes.

Every write operation must involve the parity disk however, which means there is a good chance of a bottleneck.

### RAID Level 5

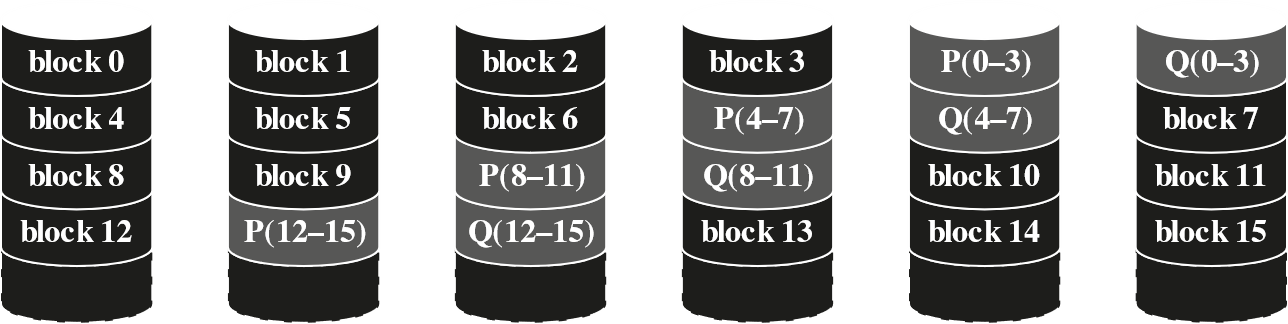
RAID 5 solves the problem of the bottleneck by distributing the parity strips across all the disks like this:



Thus, depending on which strip is written onto, a different disk must be accessed to update the parity strip, which removes the pressure from a single disk that was present in RAID 4.

### RAID Level 6

In RAID 6, two different parity calculations are carried out and stored in separate blocks on different disks. Thus, for user data requiring disks, RAID 6 consists of disks.



One of the parity calculations is the same as the ones used in RAID 4 and 5, but the other is an independent data check algorithm. This makes it possible to restore data even if two disks contain failures. RAID 6 provides extremely high data availability. 3 separate disks would have to fail before repairs are made for data to be lost. However, RAID 6 involves a high write penalty since every write affects two parity blocks. Read performance for RAID 6 is comparable to RAID 5.

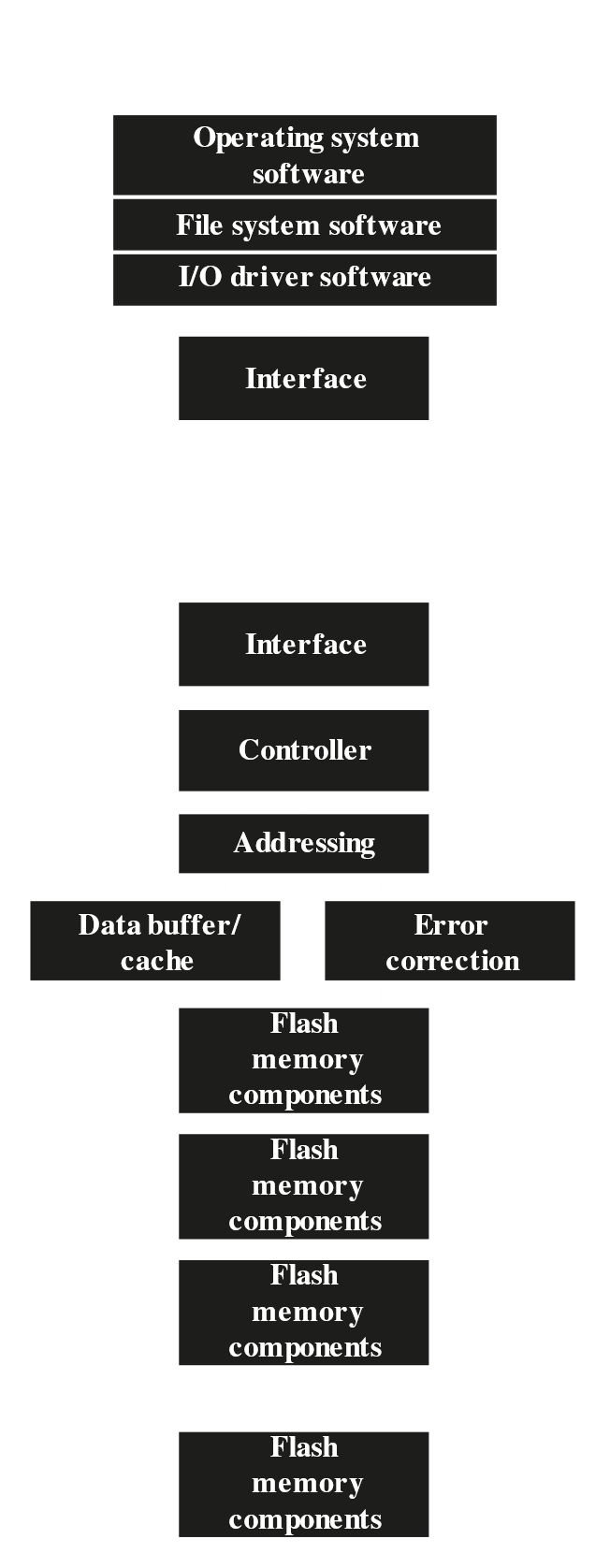
## 6.3 Solid State Drives

Solid State Drives (SSDs) are beginning to complement or even replace hard disk drives (HDDs), both as internal and external secondary memory. The term solid state refers to electronic circuitry built with semiconductors. An SSD is a memory device made with solid state components. SSDs on the market and coming on line use NAND flash memory.

### SSD Compared to HDD

As the cost of flash based SSDs has dropped and the performance and bit density has increased, SSDs have become increasingly competitive with HDDs. SSDs have significantly increased I/O performance, are less susceptible to physical shock and vibration, have a longer lifespan since they are not susceptible to mechanical wear, have lower power consumption, are quieter and cooler, requiring less energy, and have lower access times and latency rates. The disadvantages are that HDDs have much higher capacity and have lower cost per bit.

### SSD Organization



On the host system, the OS invokes the file system software to access data on the disk. The file system invokes I/O driver software, which provides the host access to the SSD. The interface component refers to the physical and electrical interface between the host processor and the SSD peripheral device. If the device is internal, PICe is commonly used, while USB is used to external devices.

The SSD also contains the following components:

* Controller to provide the device with level interfacing and firmware execution
* Addressing, which is the logic that performs selection function across the flash memory components
* Data buffer/cache, where high speed RAM memory components are used for speed matching and to increase data throughput
* Error correction for error detection and correction
* Flash memory components, which are individual NAND flash chips

### Practical Issues

There are two practical issues faced by SSDs that are not faced by HDDs. First, SSD performance tends to slow down as the device is used. Files are stored on disk as a set of pages, typically 4KB in length. These pages are not typically stored contiguously on the disk. Flash memory however, is accessed as blocks, which are typically of 512KB. To write data, the entire block must be placed in a RAM buffer, the appropriate page in the RAM buffer updated, the block of flash memory erased and the entire block from the buffer written back onto flash memory. Over time, because of how virtual memory works, files become fragmented with pages scattered over multiple blocks. More fragmentation means writing a new file involves more blocks. Thus, the writing process becomes slower. Manufacturers have tried different techniques to overcome the issue, such as setting aside a large portion of the SSD as extra space for write operations, called over-provisioning, then erasing inactive pages during idle time used to defragment the disk. Another technique is called the TRIM command, which allows an operating system to inform an SSD which blocks of data are no longer considered in use and can be wiped internally.

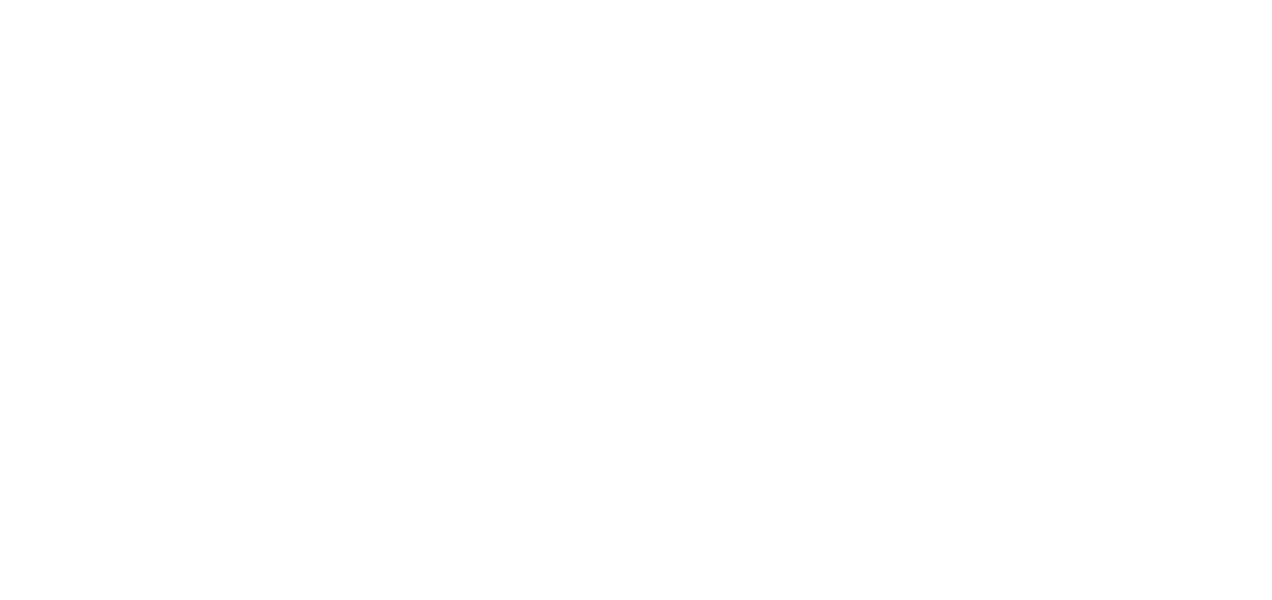
A second issue is that flash memory becomes unusable after a certain number of writes, typically 100,000. As flash cells are stressed, they lose their ability to record and retain values. Techniques for prolonging the life of an SSD drive include front-ending the flash with a cache to delay and group write operations, using wear-levelling algorithms that evenly distribute writes across a block of cells, and sophisticated bad-block management techniques. In addition, vendors are deploying SSDs in RAID configuration to further reduce the probability of data loss. Most flash devices are also capable of estimating their own remaining lifetimes, so systems can anticipate failure and take pre-emptive action.

## 6.4 Optical Memory

|  |  |  |
| --- | --- | --- |
| CD | Compact Disk | Non-erasable disk that stores digitized audio information. The standard system uses 12-cm disks and can record more than 60 minutes of uninterrupted audio. |
| CD-ROM | Compact Disk Read-Only Memory | Non-erasable disk for storing computer data. The standard system uses 12-cm disks and can hold more than 650 MB. |
| CD-R | CD Recordable | Similar to CD-ROM. The user can only write to the disk once. |
| CD-RW | CD Rewritable | Similar to CD-ROM. The user can erase and rewrite to the disk multiple times. |
| DVD | Digital Versatile Disk | A technology for producing digitized, compressed representation of video information and large volumes of other digital data. Both 8 and 12 cm diameters are used, with double sided disks having capacities up to 17 GB. The basic DVD is read-only (DVD-ROM). |
| DVD-R | DVD Recordable | Similar to DVD-ROM. The user can write to the disk only once. Only one-sided disks can be used. |
| DVD-RW | DVD Rewritable | Similar to DVD-ROM. The user can erase and rewrite to the disk multiple times. Only one-sided disks can be used. |
| Blu-ray DVD | High Definition Video Disk | Provides considerably greater data storage density than DVD, using a 405 nm (blue-violet) laser. A single layer on a single side can store 25 GB. |

### Compact Disks

#### CD and CD-ROM

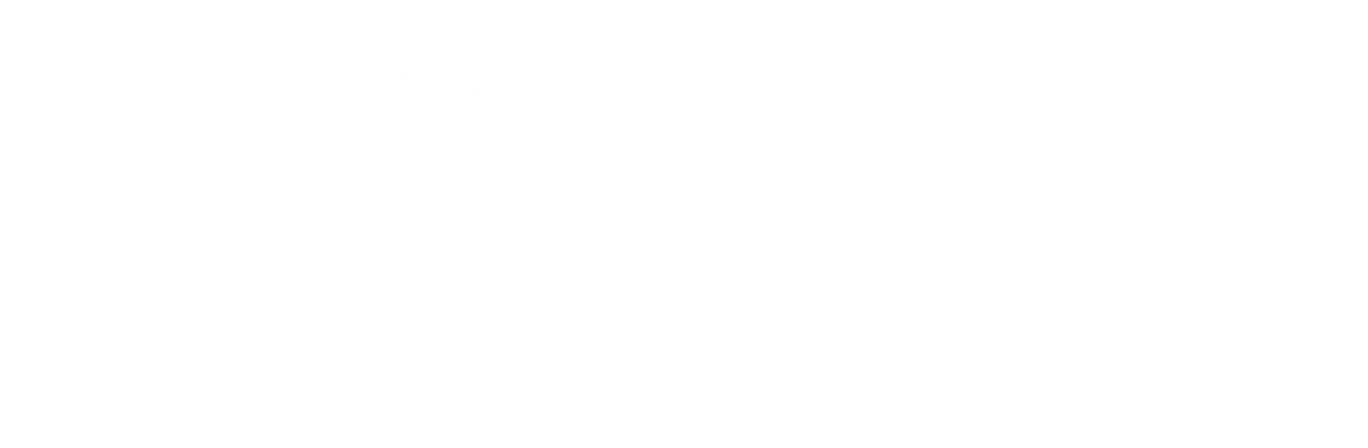


The first optical disk products where the compact disk (CD) and the compact disk read-only memory (CD-ROM). Both were similar products, made of a resin like polycarbonate, but CD-ROMs are more rugged and have error-correction properties to ensure data is properly transferred from the disk to the computer.

Digitally recorded information is imprinted as a series of pits on the surface of the polycarbonate with the help of a finely focused, high-intensity laser to create a master disk. This master is then used to stamp out copies onto polycarbonate. The pitted surface is then coated with a highly reflective material like aluminium or gold, protected from dust and scratches with a layer of clear acrylic, and a label silkscreened onto the acrylic.

Information is retrieved using a low-powered laser housed in an optical-disk player, or drive unit. The laser shines through the clear polycarbonate while a motor spins the disk past it. The intensity of the reflected light changes when it falls on a pit, which has a rough surface and thus scatters light, reflecting a less intense light back to the source. The lands on the other hand, are smooth, and thus reflect back the light at a higher intensity. A photosensor detects these changes and converts them into digital signals. The sensor tests the surface at a regular interval. The beginning and end of a pit represents a 1, while no changes in elevation represent 0s.

To achieve greater capacity, instead of using things like MZR like in magnetic disks, CDs and CD-ROMs contain a single spiral track that begins near the centre and spirals out to the outer edge of the disks. The sectors on the outside of the disk are the same length as those on the inside, which means information is packed evenly across the disk in segments of the same size that are being scanned at the same rate by rotating the disk at a variable speed. The pits are read by the laser at a constant linear velocity (CLV), with the disk rotating more slowly for access near the outer edge than for access near the centre. Thus, the capacity of a track and the rotational delays both increase for positions nearer to the outer edge of the disk.



Data on the CD-ROM is organized as a sequence of blocks, as shown above. The Sync field identifies the beginning of a block. It consists of a byte of all 0s, 10 bytes of all 1s, and a byte of all 0s. The Header contains the block address and the mode byte. Mode 0 specifies a blank data field, mode 1 specifies the use of an error-correcting code and 2048 bytes of data and mode 2 specifies 2336 bytes of user data with no error-correcting code. Data is simply the user data. Auxiliary contains additional user data in mode 2, and a 288-byte error correcting code in mode 1.

Using CLV makes random access difficult. Locating a specific address would need the head to be moved, the rotation speed to be adjusted and making minor adjustments to find and access the specific sector.

CD-ROM is best for the distribution of large amounts of data to a large number of users. Since it is expensive to initially write on CD-ROMs, it is not good for individualized applications. Compared to traditional magnetic disks, advantages of the CD-ROM include mass, inexpensive replication, and the fact that the disk itself is removable and can thus be used for archival storage. Disadvantages include the fact that it is read-only and cannot be updated and has an access time much longer than that of a magnetic disk.

#### CD-R

For applications that only need one or a small number of copies, the write-once, read-many CD called CD Recordable (CD-R) is used. A disk is prepared in such a way that it can be subsequently written once with a laser beam of modest intensity. Thus, in a somewhat more expensive fashion, the user can write once and read the disk as well. Unlike CDs and CR-ROMs, the medium of a CD-R includes a dye layer that is used to change reflectivity and is activated with a high-intensity laser. CD-R disks can be read on CD-R or CD-ROM drives. CD-Rs are attractive for archival storage.

#### CD-RW

CD-Rewritable, or CD-RW, can be repeatedly written and overwritten. The only pure optical approach to achieve this that has proven attractive is called phase change. The phase change disk uses a material with two significantly different reflectivities in two different phase states. The amorphous state has molecules randomly oriented and reflects light poorly, while the crystalline state has a smooth surface and reflects light well. A beam of laser light can change the material from one phase to another. The primary disadvantage of phase change optical disks is that the material eventually and permanently loses its desirable properties. Current materials can be used between 500,000 and 1,000,000 erase cycles. CD-RW competes with magnetic disks as secondary storage, and have the advantage of having less severe engineering tolerances, which results in higher reliability and a longer life.

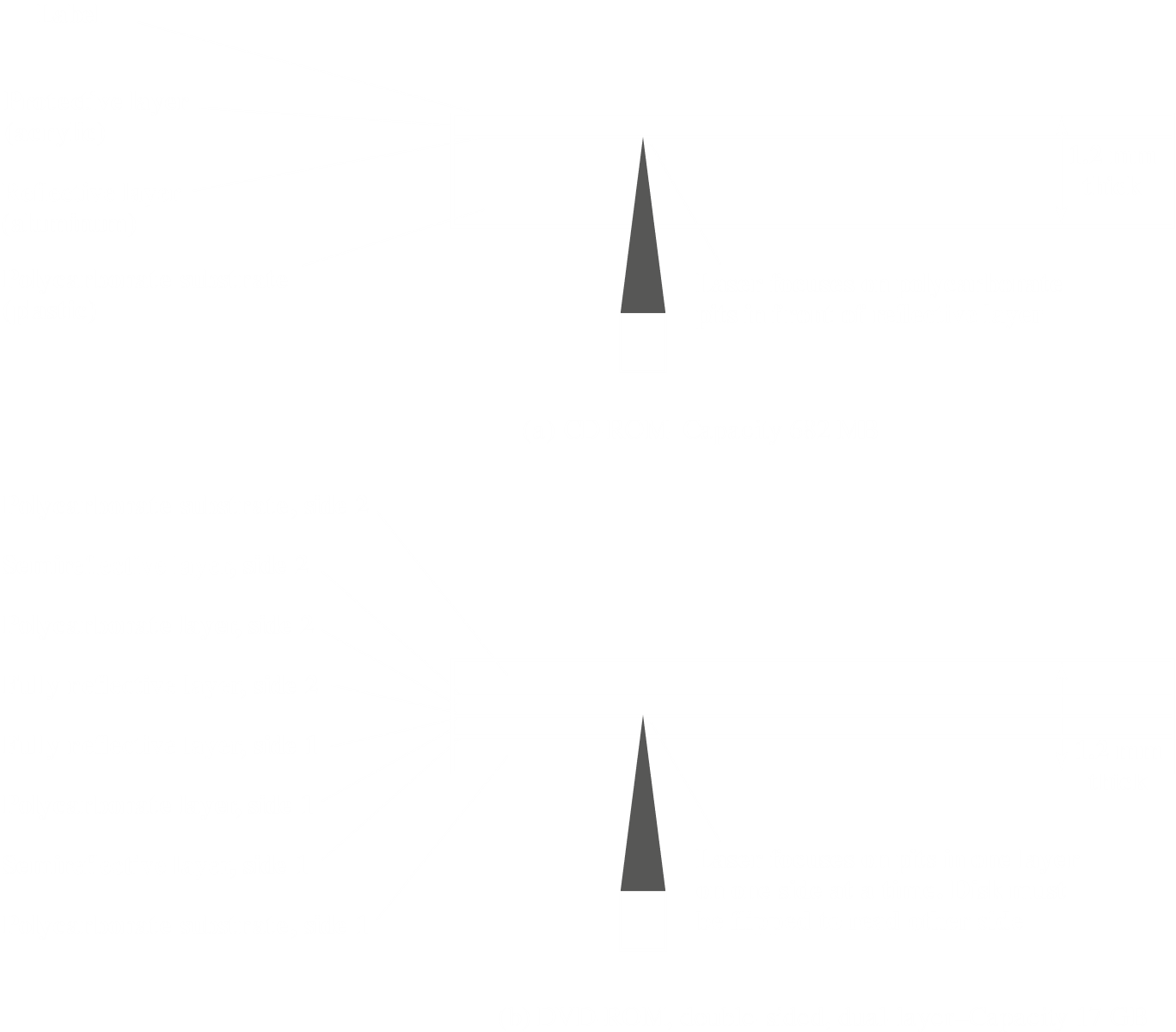
### Digital Versatile Disk

#### DVD

Digital Versatile Disks (DVDs) were an acceptable replacement for the analogue VHS video tapes. DVDs replaced both VCRs and CDs. They brought videos into the digital age, since they could deliver movies with impressive quality and could be randomly accessed, just like CDs. They can hold 7 times as much data as CDs, which made PC games more realistic and gave educational software the ability to incorporate more video. The developments lead to a new crest of traffic over the internet, as the material was incorporated into websites. Like CDs, DVDs also come in writeable and read-only versions.

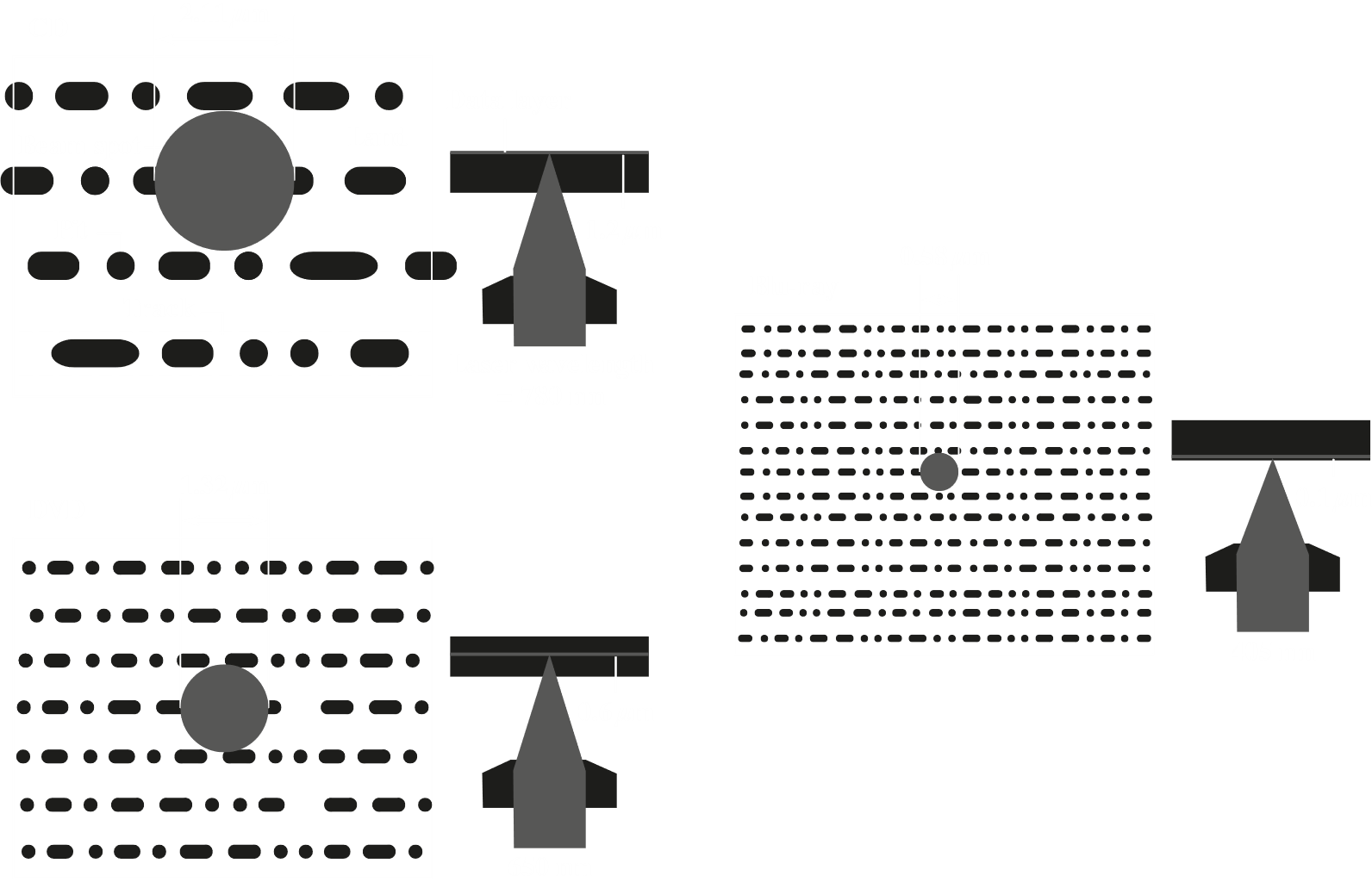
The greater storage capacity of DVDs is due to 3 main factors:

* Bits are packed more closely, with spacing between loops of a spiral being at and minimum distance between pits at , compared to the and respectively for CDs. This is due to the use of a laser with a shorter wavelength in DVDs. The result is a seven-fold increase in capacity, to about 4.7 GB.
* DVDs use a second layer of pits and lands on top of the first layer, with a semi-reflective layer on top of the reflective layer. By adjusting focus, the lasers can read each layer separately. The lower reflectivity of the second layer limits is capacity a little, but the overall effect is still to nearly double the capacity to about 8.5 GB.
* DVD-ROMs can be two sided, as opposed to the one-sided nature of CDs, bringing the total capacity up to 17 GB.



#### Blu-Ray Disks

High-definition optical disks have even shorter wavelength lasers, in the blue-violet range, which allows for significantly greater storage capacities. This also results in smaller data pits. Although a competing format, HD DVD, also competed initially, it could only store 15 GB of data on a single layer on a single side. Blu-ray positions the data layer closer to the laser, enabling tighter focus and less distortion, resulting in smaller pits and tracks. Blu-ray can store 25 GB on a single layer, and thus won in the market. Three versions are available for Blu-ray, read only (BD-ROM), recordable once (BD-R) and rerecordable (BD-RE).

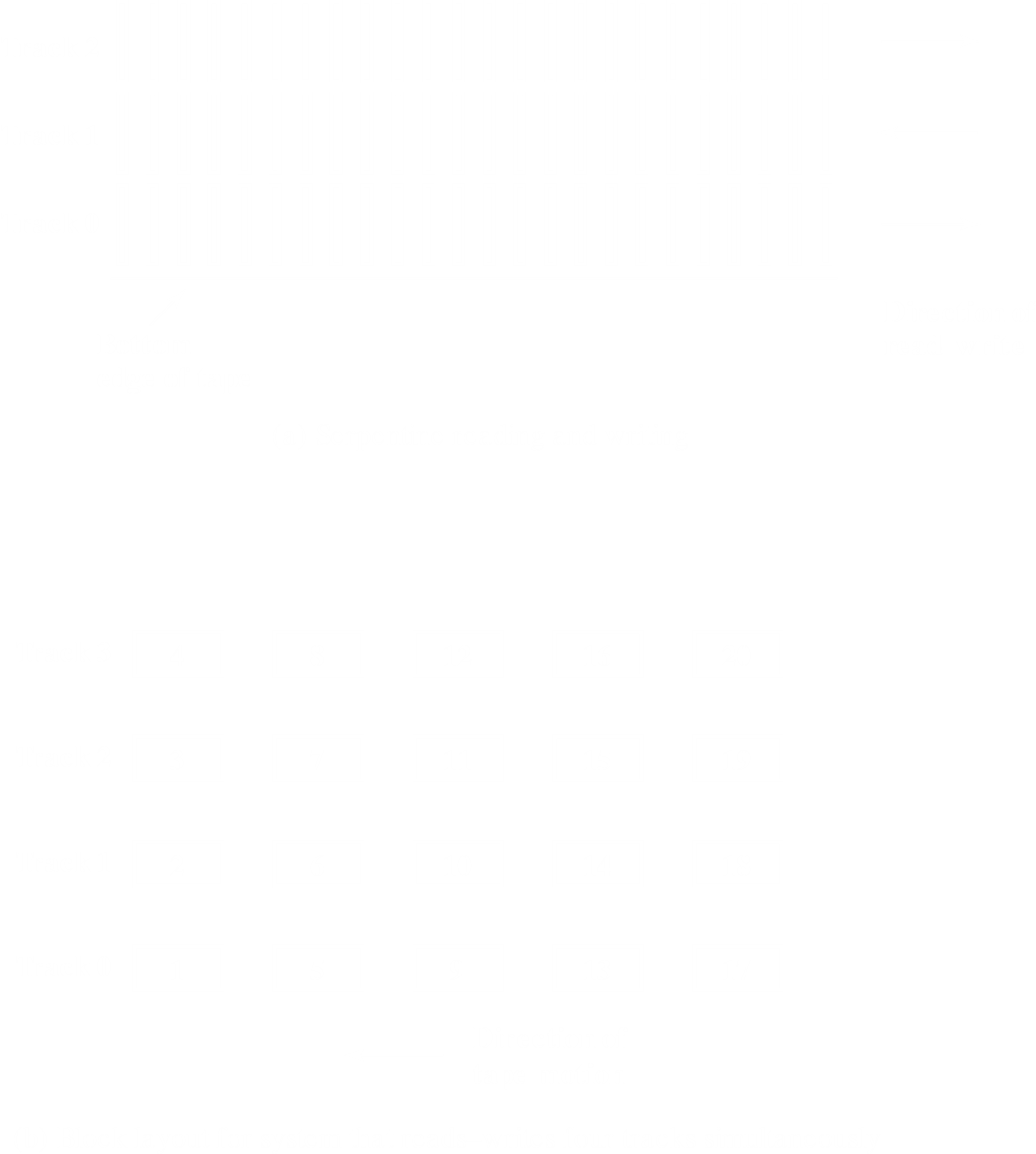


## 6.5 Magnetic Tape

Tape systems use the same reading and recording techniques as disk systems, but use a flexible polyester tape coated in magnetizable material as a medium. The coating may contain particles of pure metal in special binders or vapor-plated metal films. The tape widths vary from 0.38 cm to 1.27 cm. They used to be packaged as open reels that had to be threaded through a second spindle for use, but today almost all tapes are housed in cartridges. They were the first kind of secondary memory, and are still widely used as the lowest-cost slowest-speed member of the memory hierarchy. The dominant tape technology today is a cartridge system called linear tape-open (LTO), developed in the late 1990’s as an open-source alternative to the various proprietary systems on the market.

Data on tapes is placed on a number of parallel tracks running length-wise. Typically, there are 9 tracks, with one being used for parity bits, but later systems also used 18 or 36 tracks corresponding to a digital word size. The recording of data in this fashion is called parallel recording. In contrast, most modern systems store data as a sequence of bits along each track, called serial recording. Data on tapes is read and written in contiguous blocks, called physical records, with block being separated by gaps called inter-record gaps.

The typical recording medium in serial tapes is called serpentine recording. The first set of bits is recorded along the whole length of the first track, the second set is recorded after that, starting in the opposite direction and so on. It is possible to increase speed by reading/writing multiple tracks simultaneously, typically two to eight tracks.



A tape drive is a sequential-access device, meaning it needs to go through records one at a time in a fixed serial. If an older record needs to be read, the tape needs to be rewound. Unlike disks, the tape is only in motion during read or write operations. By contrast, disk drives are direct-access devices. They do not need to read all the sectors sequentially, but must only wait for the intervening sectors on the same track and can make successive accesses to any track.