

Architektura počítačů

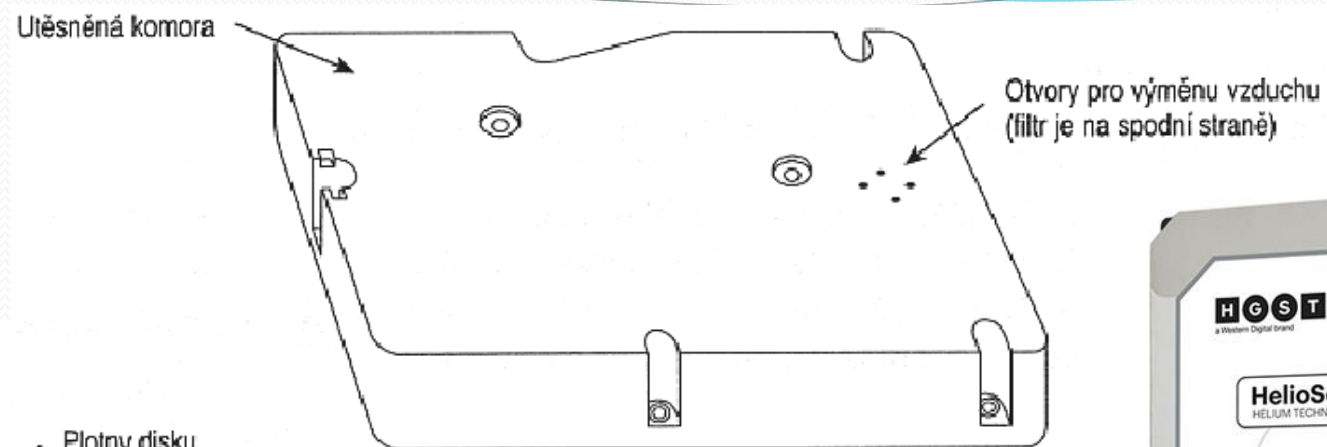
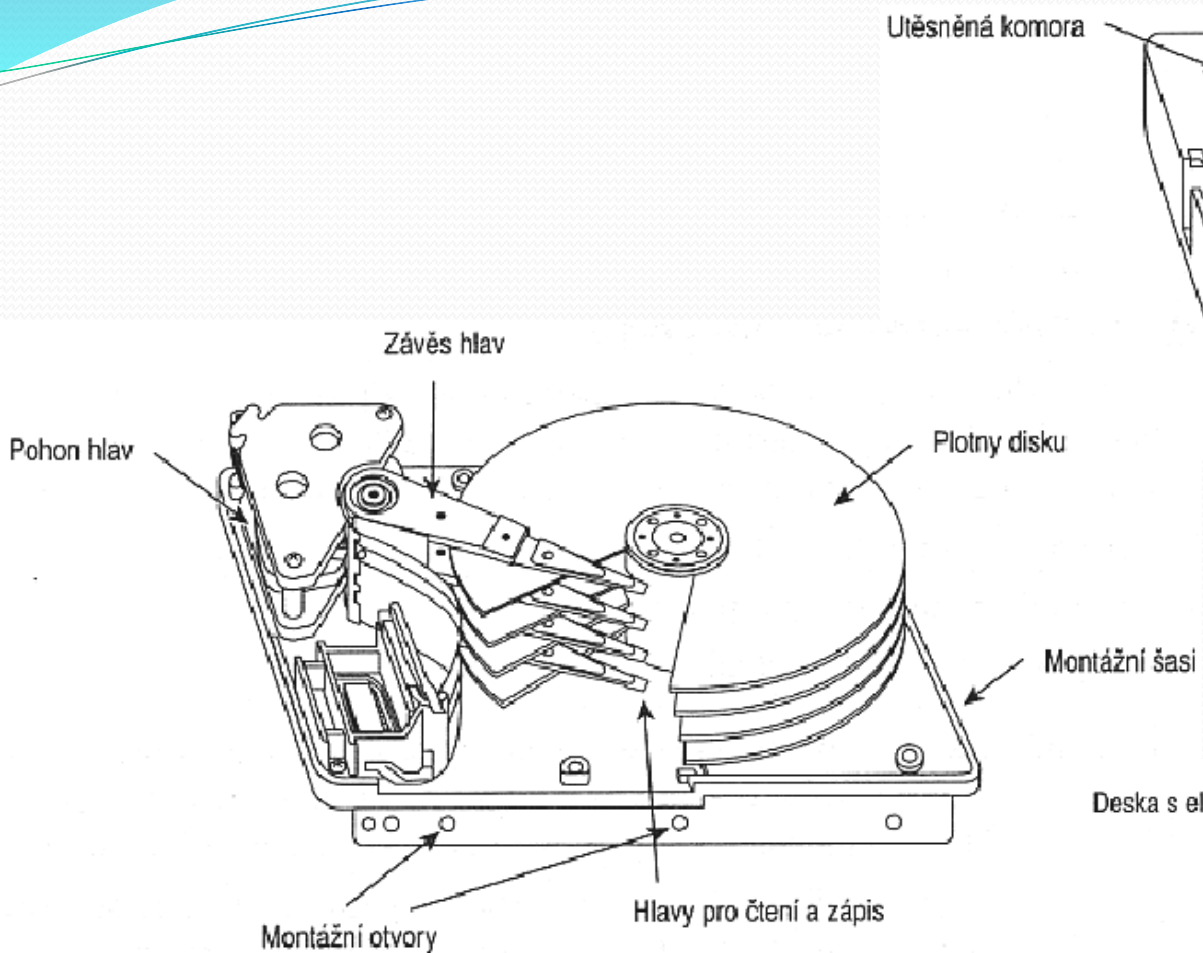
07

Pevné disky

- Média pro uchování dat
- Kapacita se udává v TB (2^{40} nebo 10^{12} B)
- Magnetický záznam dat na rotující plotny
- Používané typy se dělí podle
 - Fyzické velikosti ($3\frac{1}{2}$ palce, $2\frac{1}{2}$ palce, 1 palec)
 - Typ rozhraní (ATA, SCSI)
- Rychlosti otáčení (3600 rpm – 15000 rpm)
- Přístupová doba (4 -25 ms)

Základní součásti pevného disku

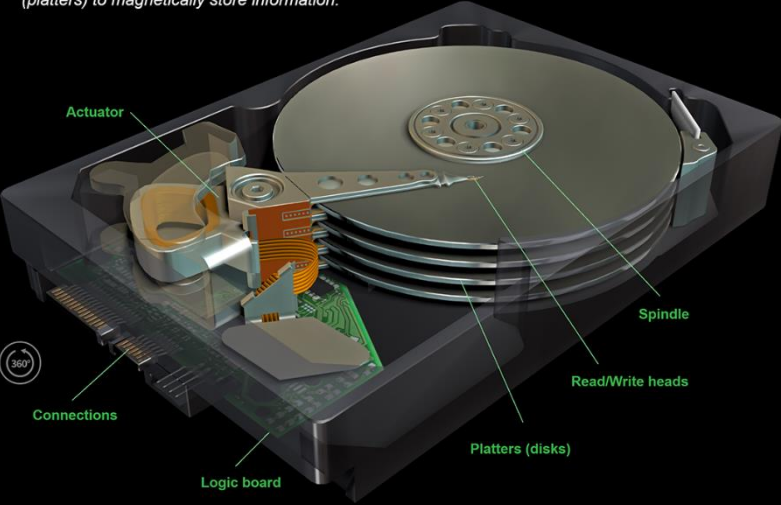
- Plotny disku
- Hlavy pro čtení a zápis
- Pohon hlav
- Pohon ploten disku
- Vzduchové filtry
- Řídící deska (deska s elektronikou)
- Kabely a konektory
- Případné konfigurační prvky (např. propojky a přepínače)



How Hard Disk Drives Work

Presented by SEAGATE.COM
Designed & research by ANIMAGRAFFS

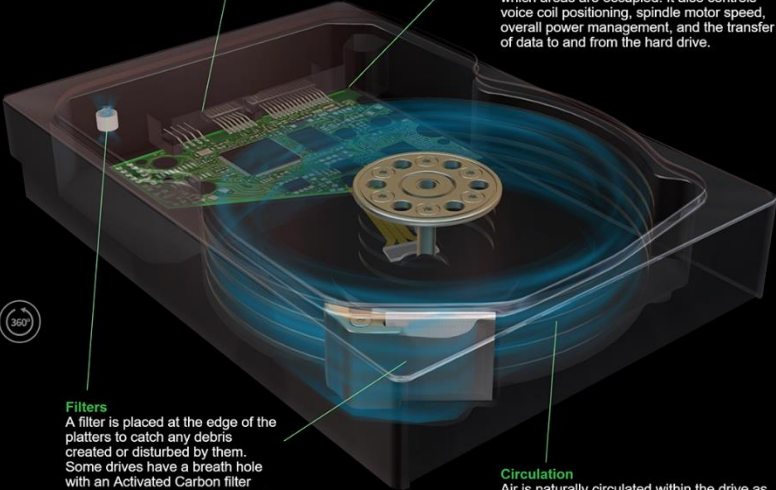
Hard Disk Drives (HDDs) use spinning disks (platters) to magnetically store information.



Case

Connections
Hard drives have separate connections for the power and data cables, and a Jumper Block. The pins on the jumper block can be connected to slow data transfer rate to be compatible with older, slower hard drives.

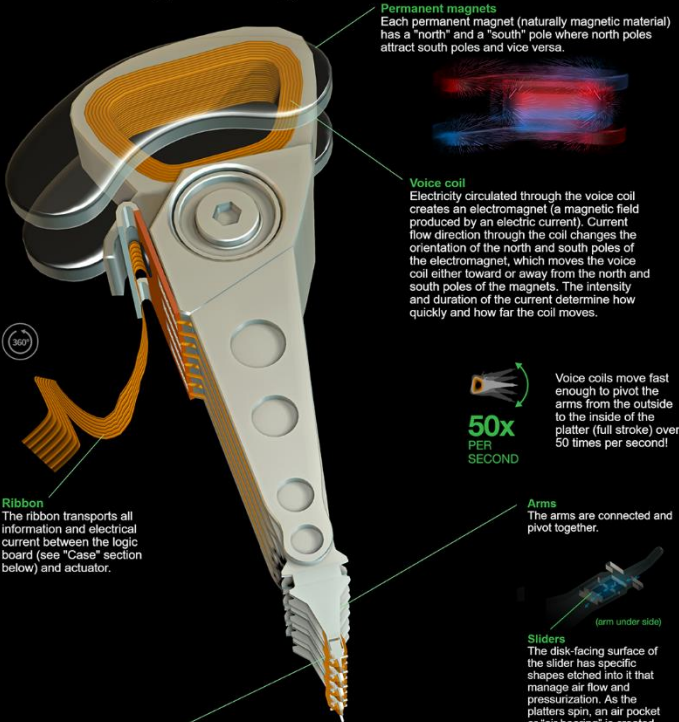
Logic board
The logic board controls the read/write heads using a built-in map of the platters to determine which areas are available and which areas are occupied. It also controls voice coil positioning, spindle motor speed, overall power management, and the transfer of data to and from the hard drive.



Filters
A filter is placed at the edge of the platters to catch any debris created or disturbed by them. Some drives have a breath hole with an Activated Carbon filter that absorbs vapors, and prevents dust or debris from entering the drive. Newer helium drives are sealed and don't have a breath hole.

Circulation
Air is naturally circulated within the drive as the platters spin, while a breather hole allows the interior pressure to equalize with air pressure on the outside of the drive.

Reading & writing data



Permanent magnets
Each permanent magnet (naturally magnetic material) has a "north" and a "south" pole where north poles attract south poles and vice versa.

Voice coil
Electricity circulated through the voice coil creates an electromagnet (a magnetic field produced by an electric current). Current flow direction through the coil changes the orientation of the north and south poles of the electromagnet, which moves the voice coil either toward or away from the north and south poles of the magnets. The intensity and duration of the current determine how quickly and how far the coil moves.

50x PER SECOND
Voice coils move fast enough to pivot the arms from the outside to the inside of the platter (full stroke) over 50 times per second!

Ribbon
The ribbon transports all information and electrical current between the logic board (see "Case" section below) and actuator.

Arms
The arms are connected and pivot together.

Sliders
The disk-facing surface of the slider has specific shapes etched into it that manage air flow and pressurization. As the platters spin, an air pocket or "air bearing" is created and keeps the head ~2 nm (nanometer) from each platter – thinner than a finger print.

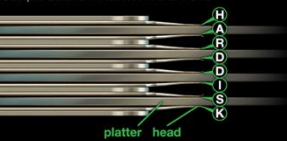
Read-write heads

Read heads
The read heads are TMR (tunneling magnetoresistance) devices, which consist of an insulator sandwiched between two sensitive magnetic materials.

Magnetic fields from the platter influence the closest magnet, causing electrons to tunnel or travel through the insulator and flip the polarity of the second magnet – thus, "reading" the platter's varying fields without disturbing them.

Magnetic domain
Electrons in atoms create magnetic fields, and the direction they "spin" determines the direction of their north and south magnetic poles. Hard disk magnetic domains usually consist of about 100,000 atoms with magnetic poles oriented in the same direction.

Sample data to write: H A R D D I S K

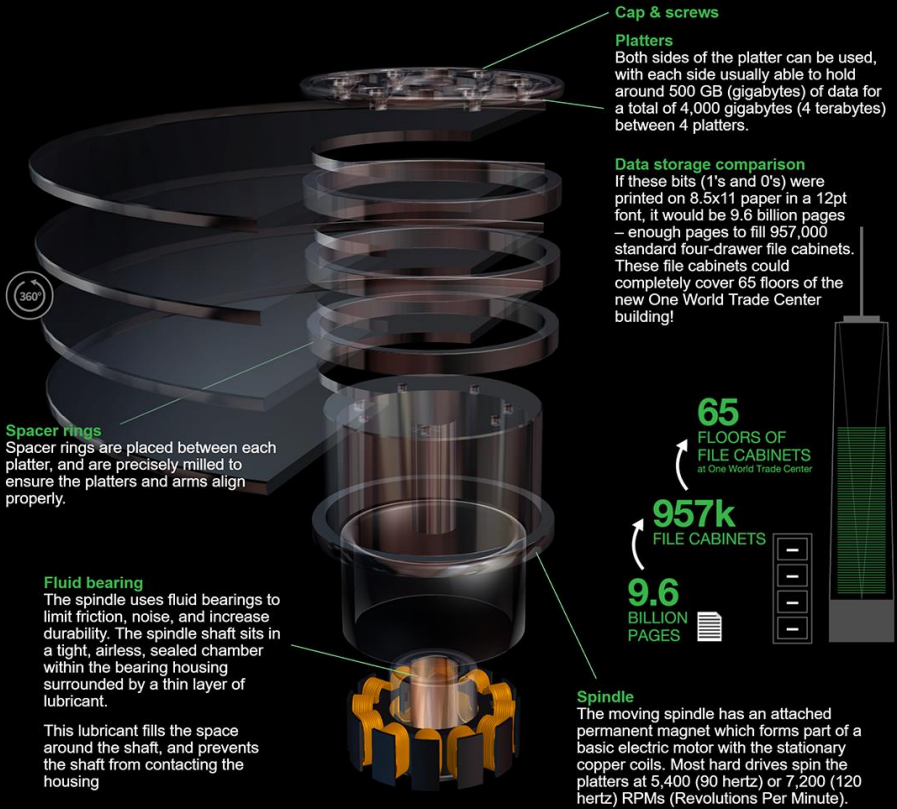


Multi-platter writing process

Data is written on both sides of each platter in the same respective location; if something 16 bits in size is written to or read from the platters, 2 bits would be on each of the 8 platter sides to account for the total amount.

For example, the phrase HARD DISK could be written on both sides of four platters as shown in the graphic to the left.

Platters & spindle



Cap & screws
Platters
Both sides of the platter can be used, with each side usually able to hold around 500 GB (gigabytes) of data for a total of 4,000 gigabytes (4 terabytes) between 4 platters.

Data storage comparison
If these bits (1's and 0's) were printed on 8.5x11 paper in a 12pt font, it would be 9.6 billion pages – enough pages to fill 957,000 standard four-drawer file cabinets. These file cabinets could completely cover 65 floors of the new One World Trade Center building!

65 FLOORS OF FILE CABINETS at One World Trade Center
957k FILE CABINETS
9.6 BILLION PAGES

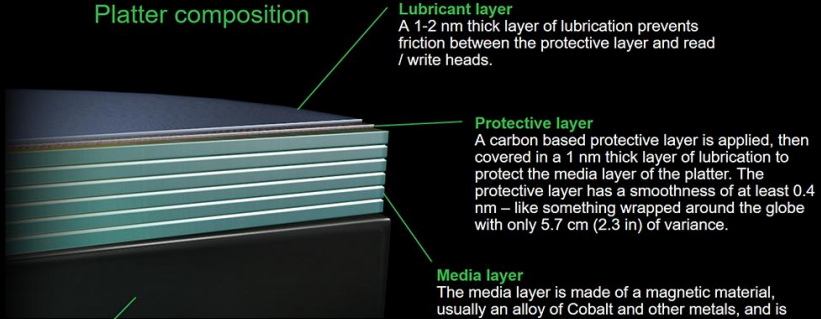
Spacer rings
Spacer rings are placed between each platter, and are precisely milled to ensure the platters and arms align properly.

Fluid bearing
The spindle uses fluid bearings to limit friction, noise, and increase durability. The spindle shaft sits in a tight, airless, sealed chamber within the bearing housing surrounded by a thin layer of lubricant.

This lubricant fills the space around the shaft, and prevents the shaft from contacting the housing

Spindle
The moving spindle has an attached permanent magnet which forms part of a basic electric motor with the stationary copper coils. Most hard drives spin the platters at 5,400 (90 hertz) or 7,200 (120 hertz) RPMs (Revolutions Per Minute).

Platter composition



Lubricant layer
A 1-2 nm thick layer of lubrication prevents friction between the protective layer and read / write heads.

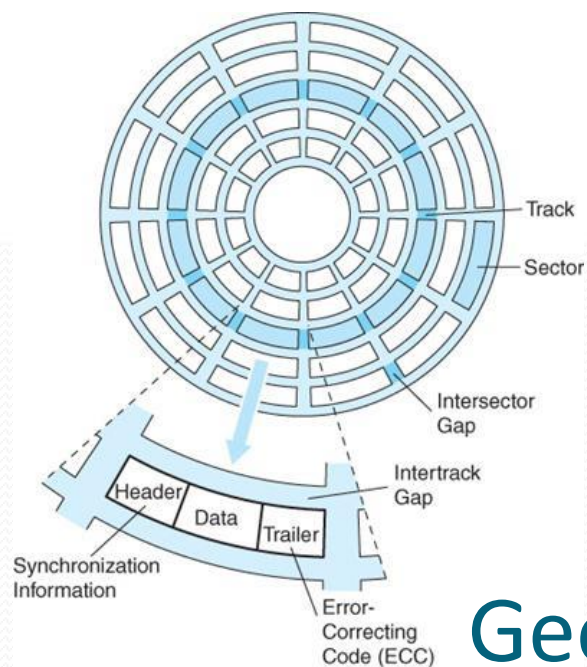
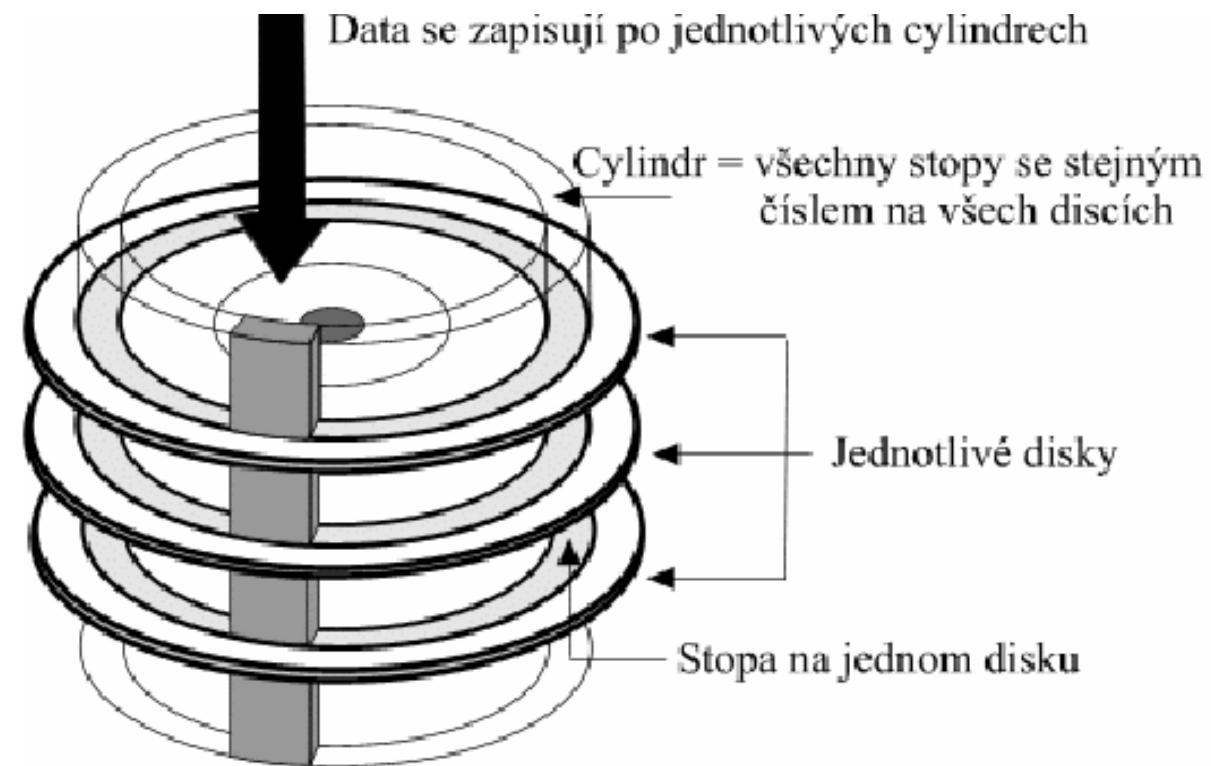
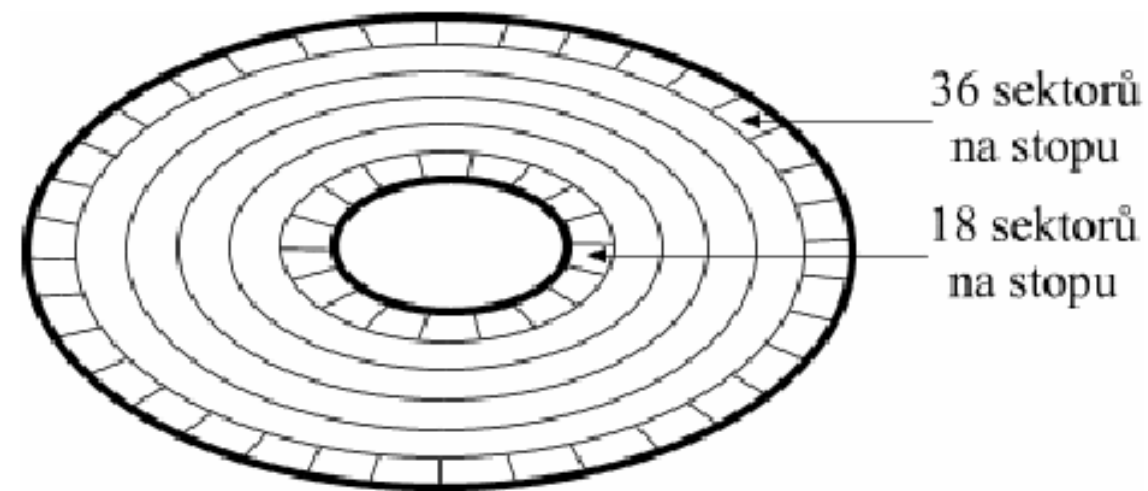
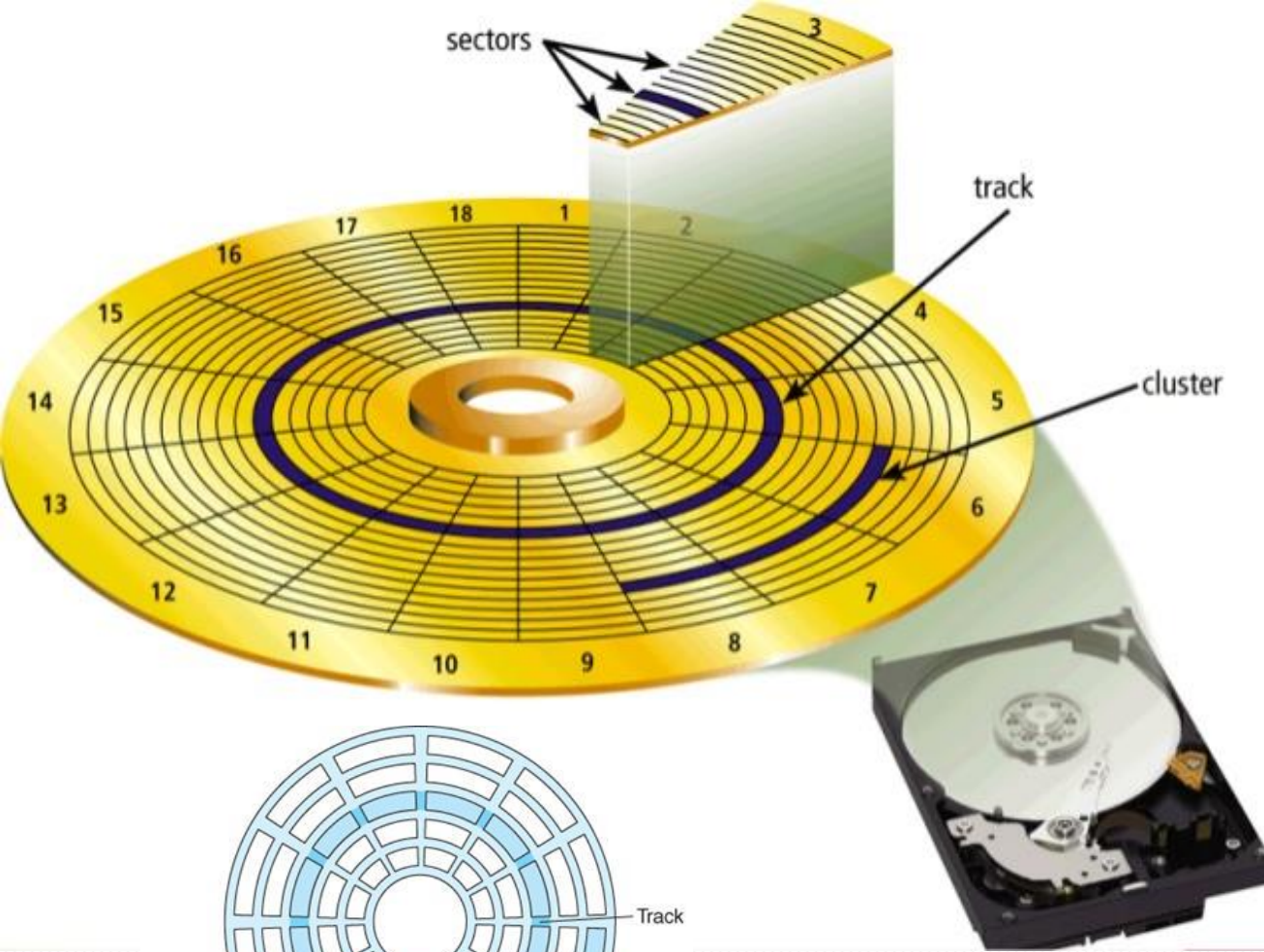
Protective layer
A carbon based protective layer is applied, then covered in a 1 nm thick layer of lubrication to protect the media layer of the platter. The protective layer has a smoothness of at least 0.4 nm – like something wrapped around the globe with only 5.7 cm (2.3 in) of variance.

Media layer
The media layer is made of a magnetic material, usually an alloy of Cobalt and other metals, and is about 100 atoms thick.

Substrate
The base of the platter is non-magnetic and is usually made from aluminum or glass (this layer would be many times thicker than those above if shown at full width).

Geometrie pevných disků

- **Stopa (track)** – soustředné kružnice na disku – číslované od kraje disku – od nuly
- **Sektor (sector)** – malá část stopy, nejmenší jednotka pro ukládání dat
 - 512 B nebo 4kiB
- **Válec (cylinder)** - množina všech stop se stejným číslem na všech površích

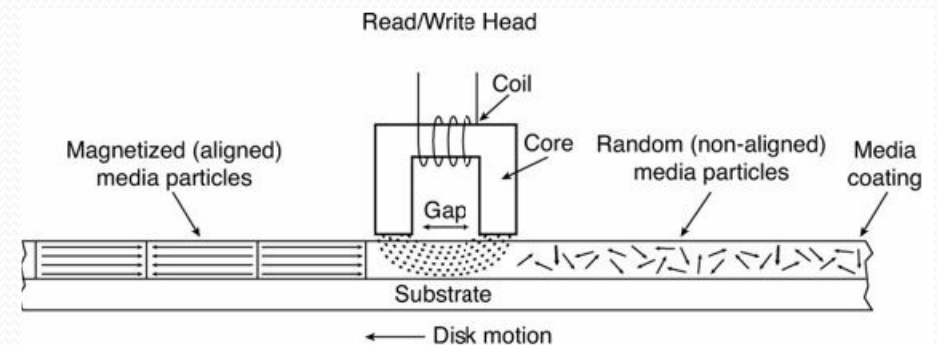
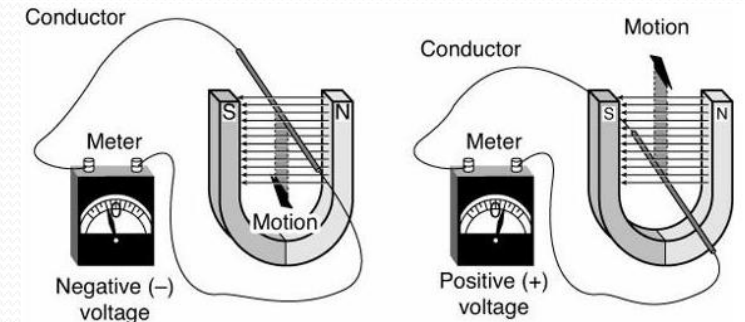


Geometrie pevných disků

Vztah mezi stopami a cylindry

Modulace dat při záznamu

- Povrch média pohybuje pod hlavou
 - Generuje se magnetický tok dané polarity
- Obrácení toku je změna polarity uspořádaných magnetických částic.
- Záznam pomocí změn magnetického toku
 - z kladného toku na záporný nebo naopak ze záporného na kladný.
 - změna se při čtení projeví jako **impuls**
- Požadavky
 - Nesmí být dlouhé posloupnosti mezer
 - Počet impulsů co nejmenší.
- Metoda kódování - konkrétní vzorec zvrátů v přechodových buňkách



Modulace dat při záznamu

- **FM modulace:** (Frequency Modulation)

Bit	Zakódování
0	PN
1	PP

Například: 101101101 se zakóduje jako PPPNPPPPPNPPPPPNPP.

Modulace dat při záznamu

- **MFM modulace:**(Modified Frequency Modulation)

Bit	Zakódování
0	PN jestliže je v řetězci 00 NN jestliže je v řetězci 10
1	NP

Např. **101100**

NPNNPNPNPNPN

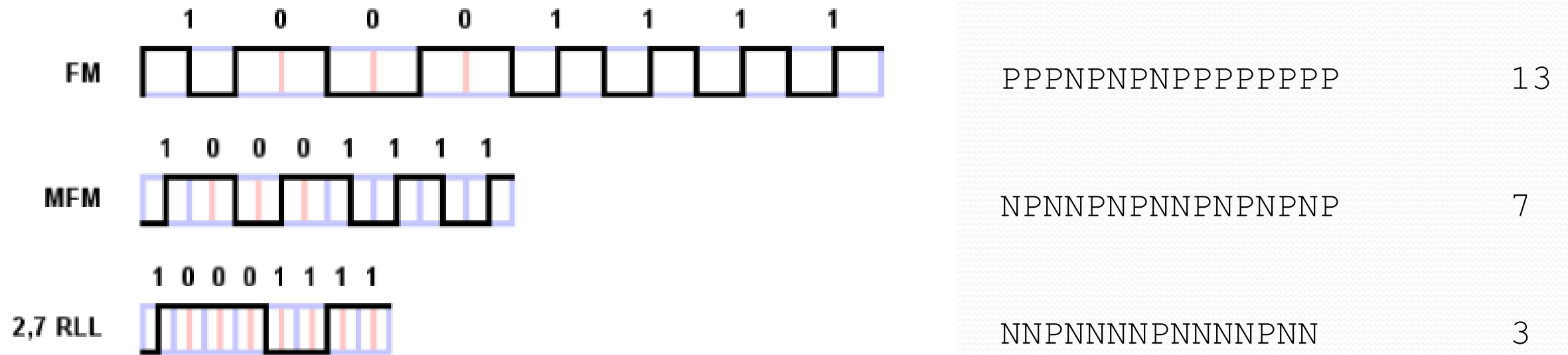
Modulace dat při záznamu

- RLL modulace (Run Length Limited)

Vzorek	RLL	Počet impulsů
00	PNNN	1
01	NPNN	1
100	NNPNNN	1
101	PNNPNN	2
1100	NNNNPNNN	1
1101	NNPNNPNN	2
111	NNNPNN	1

Vzorek	RLL	Počet impulsů
11	PNNN	1
10	NPNN	1
011	NNPNNN	1
010	PNNPNN	2
000	NNNPNN	1
0010	NNPNNPNN	2
0011	NNNNPNNN	1

Modulace dat při záznamu



Každá změna magnetického toku (impuls – log. 1) vyžaduje 2 obráceně orientované oblasti

- Moderní pevné disky používají většinou nějakou modifikaci 2,7 RLL kódování, označovanou např. **ARLL** (Advanced RLL), **ERLL** (Enhanced RLL), **PRML** (Partial Response Maximum Likelihood), **EPRML**, apod.,
- poskytují ještě větší úsporu a tím je umožněn záznam většího objemu dat na médium.

Modulace dat na optickém mediu

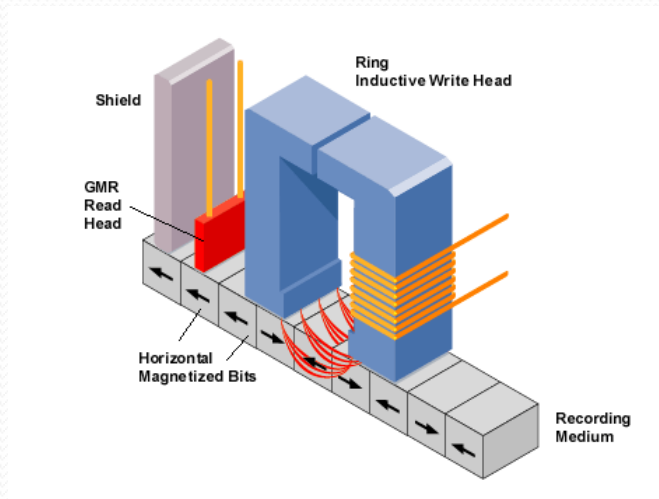
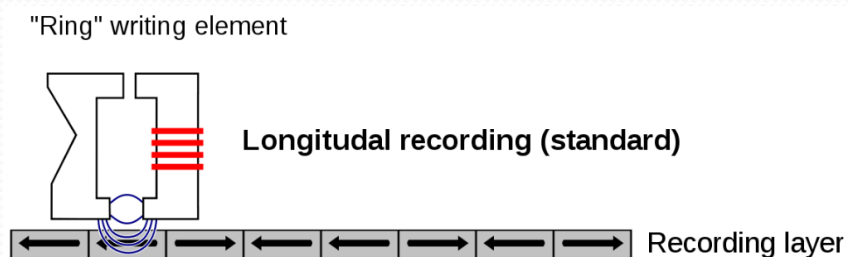
- Technologie neumožňuje zapisovat za sebou více než jednu jedničku
 - jednička vzniká jako změna stavu, tedy na přechodu land-pit, pit-land
- CD využívá se kódování osm na čtrnáct (*Eight to Fourteen Modulation – EFM*)
 - vyloučena možnost existence dvou jedniček za sebou.
- Jeden byte na CD má tedy 14 bitů + 3 oddělovací „merge“ bity
- 00000000 je překódováno na 01001000100000
00000001 je překódováno na 10000100000000

Modulace dat na optickém mediu

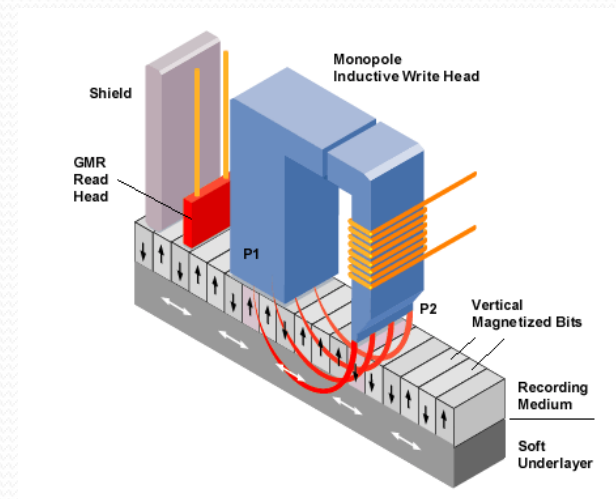
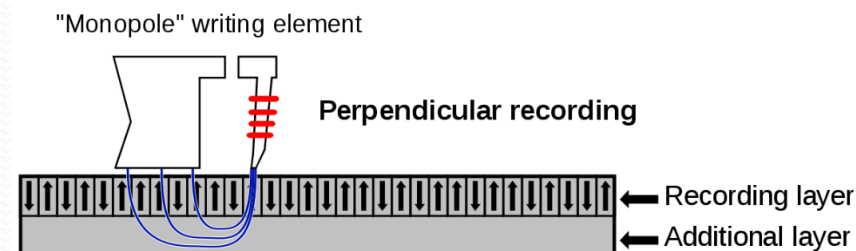
- překódování z 8bitového zobrazení 256 hodnot do 14bitového zobrazení
- kódování a dekódování je prováděno pomocí kódovací tabulky,
- slouží ke snížení počtu chyb, a to tím, že se zmenšuje počet přechodů 1 -> 0 a 0->1, a tím pádem je také třeba méně pitů.
- DVD překódování z 8bit. do 16bit.
- HD-DVD překódování z 8bit. do 12bit. ETM (Eight to Twelve Modulation)
- Blue-ray - 17PP – proměnný počet bitů

Metody navýšení hustoty zápisu

Podélný záznam

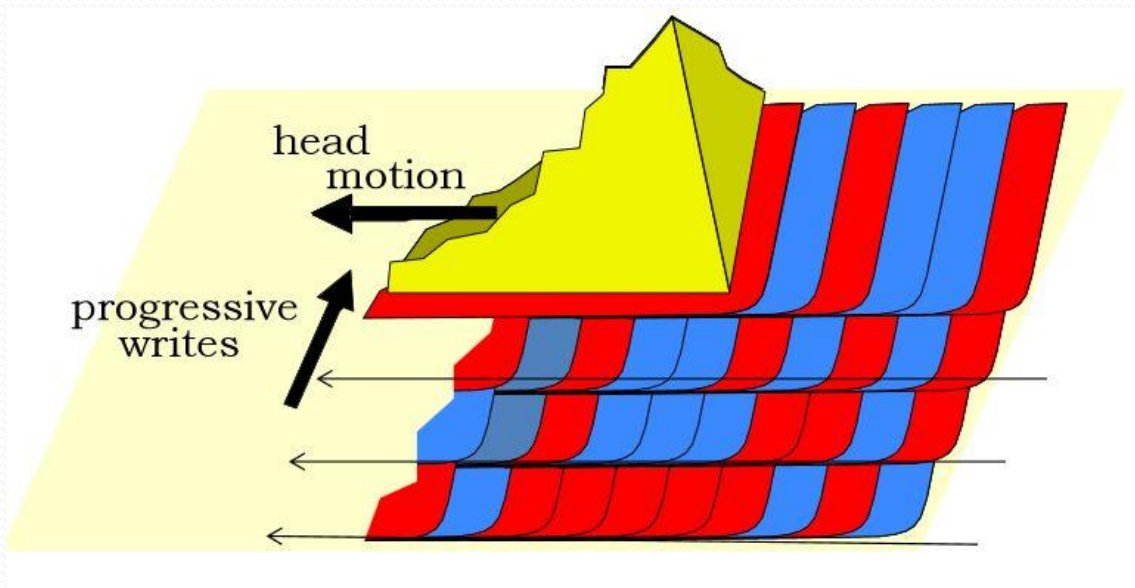


Kolmý záznam

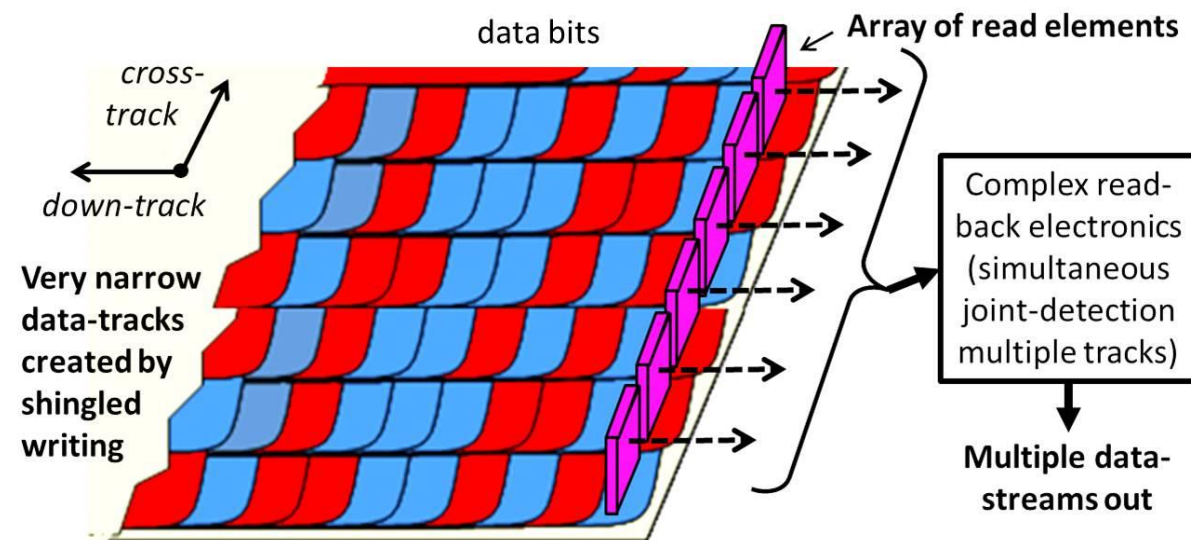


Metody navýšení hustoty zápisu

Shingled magnetic recording



Two-Dimensional Magnetic Recording (TDMR)

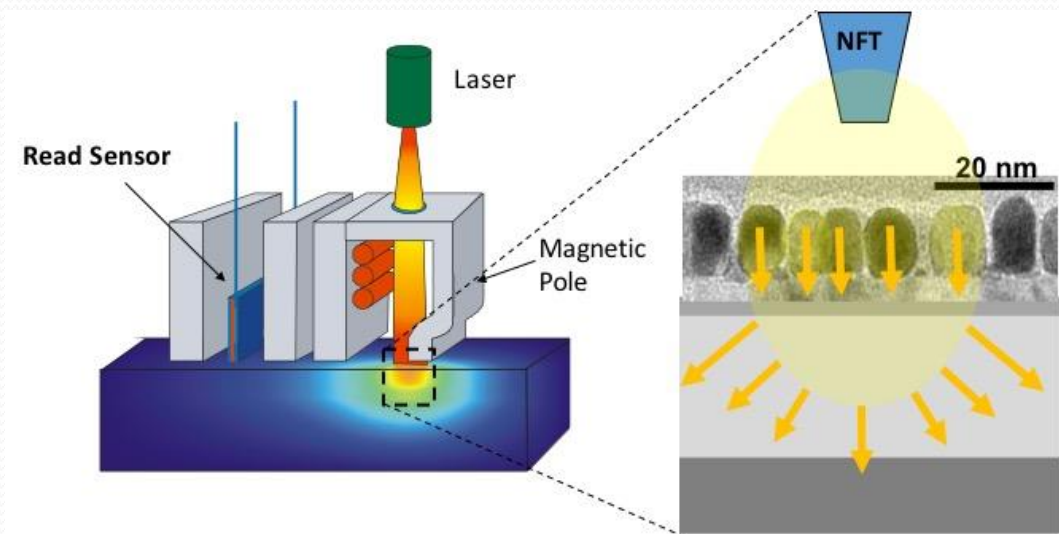


Generalized concept of TDMR as originally proposed

Metody navýšení hustoty zápisu

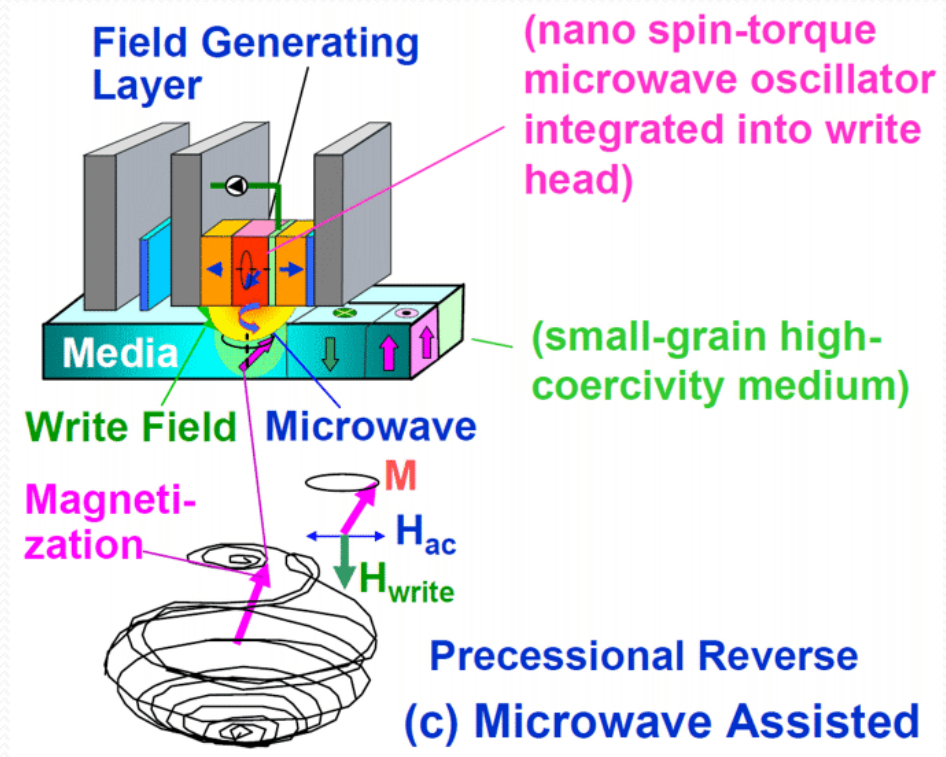
HAMR

Heat-Assisted Magnetic Recording



MAMR

Microwave-Assisted Magnetic Recording



Energy-Assisted Magnetic Recording
(Western Digital)

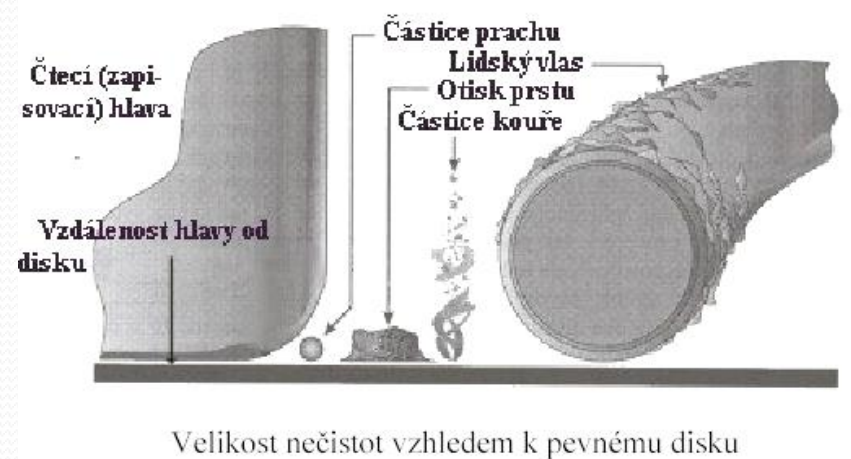
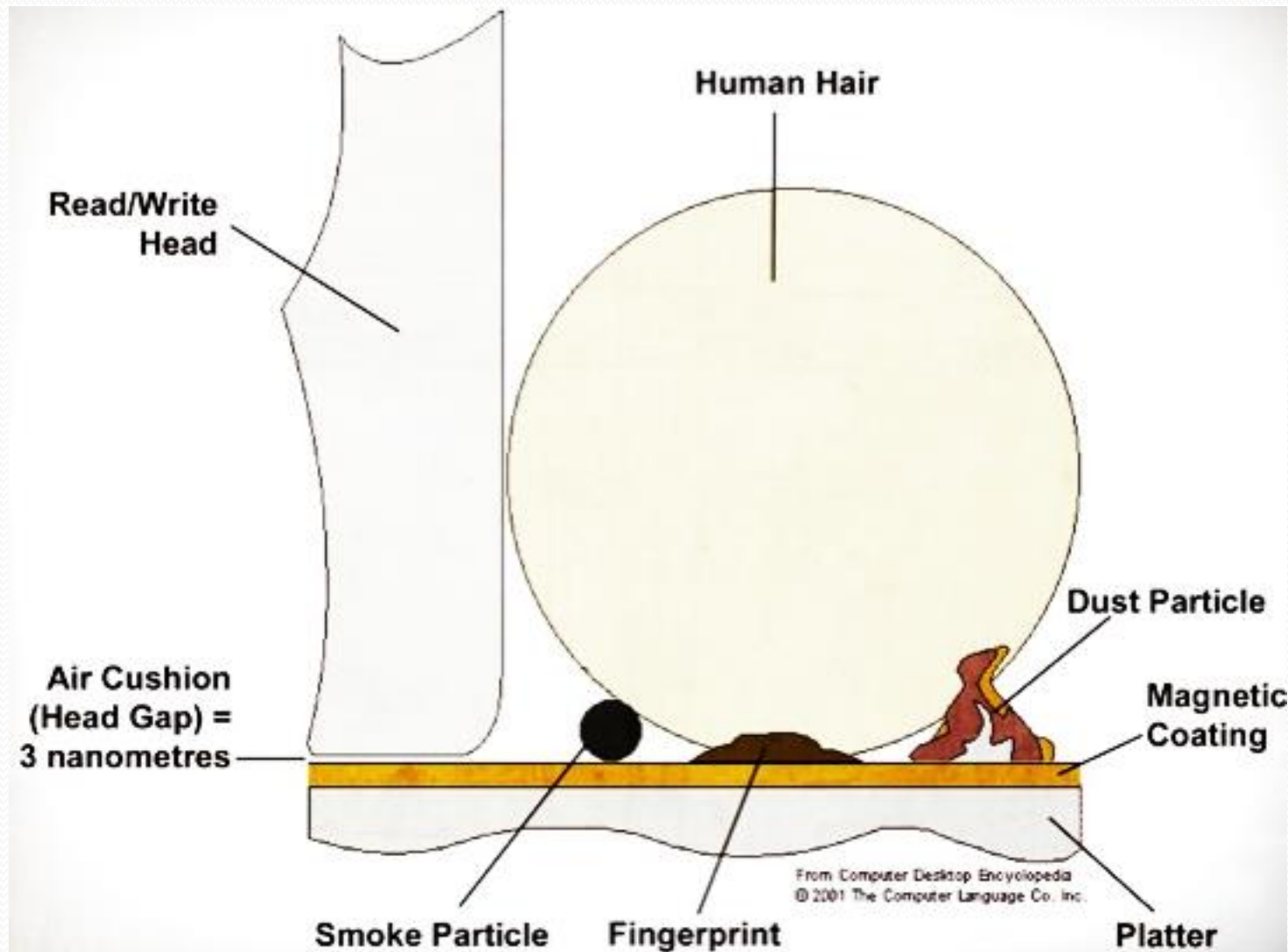
Plotny pevného disku

- Velikost disku v palcích udává velikost plotny
 - $3\frac{1}{2}$ palce
 - $2\frac{1}{2}$ palce
 - 1 palec
- Vyrábí se ze slitiny hliníku a hořčíku
 - IBM se pokoušelo o sklokeramický materiál
- Plotna je pokryta tenkou vrstvou magnetické látky
 - Oxidové médium (0.0008 mm)
 - Thin Film
 - Plátovaný TF (slitiny kobaltu) (elektrolyticky) (0.00008 mm ~ 80 nm)
 - Stříkaný TF (sloučeniny niklu a fosforu, slitiny kobaltu) (0.00003 mm ~ 30 nm)
 - Médium AFC (mezivrstva ruthenia)

Hlavy pro čtení a záznam

- Obvykle jedna hlava pro čtení a jedna hlava pro záznam na každý povrch
- Hlavy zavěšeny na pohyblivém závěsu (ramena)
- V klidu jsou hlavy zaparkovány – **automatické parkování**
 - Po roztočení disku začne proudit i vzduch
 - hlavy plují na vzduchovém polštáři nad povrchem





Lidský vlas 42-95 μm

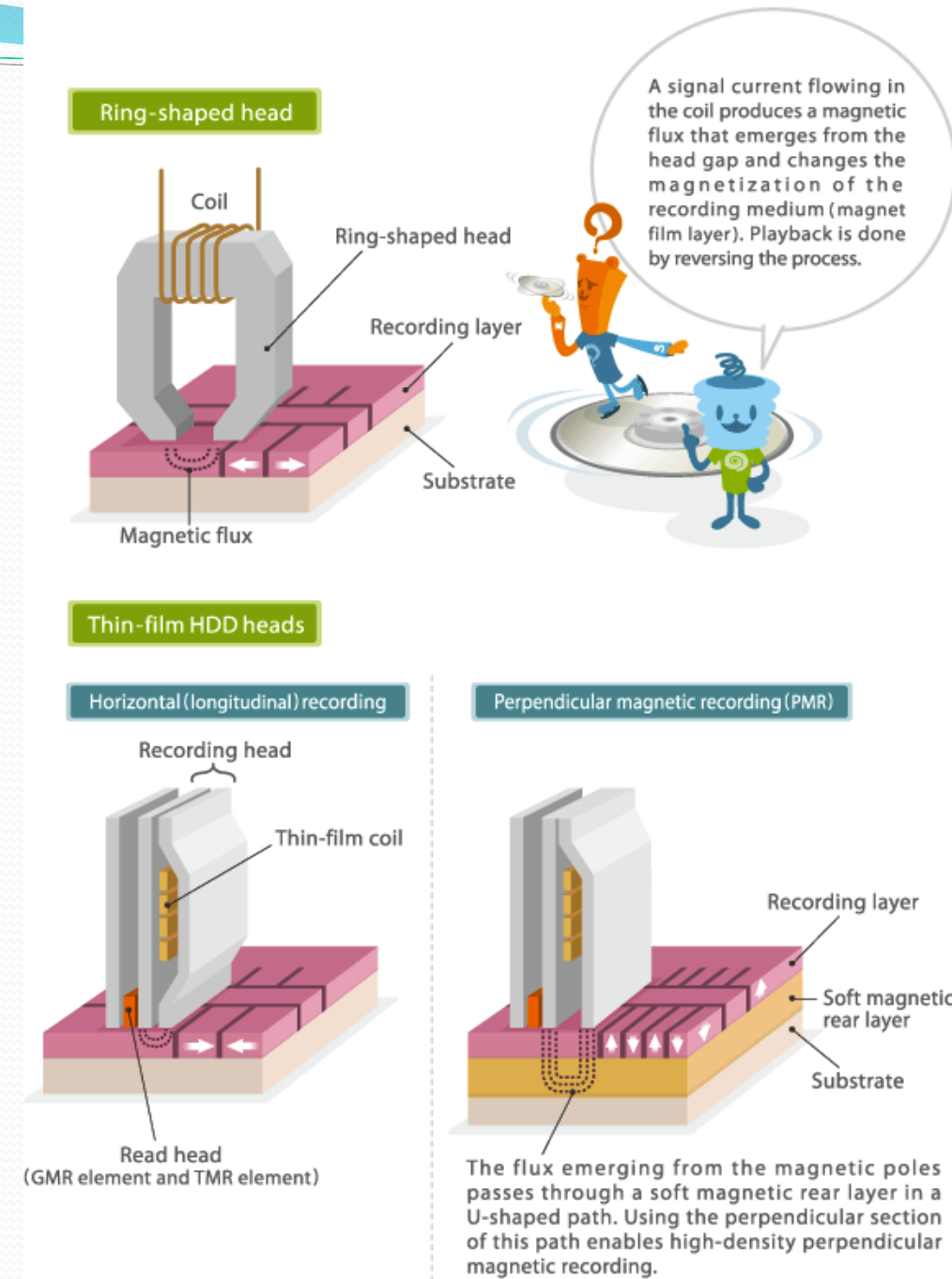
Částice prachu > 1-10 μm

Vzdálenost hlavy od plotny
~ 3 nm

Konstrukce hlav

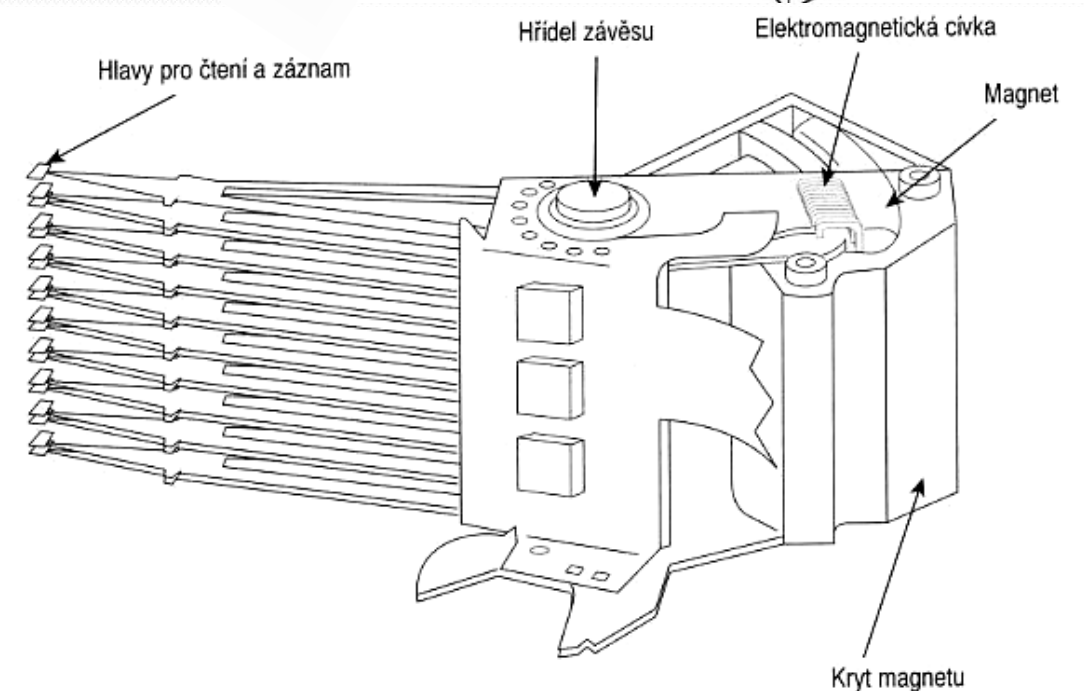
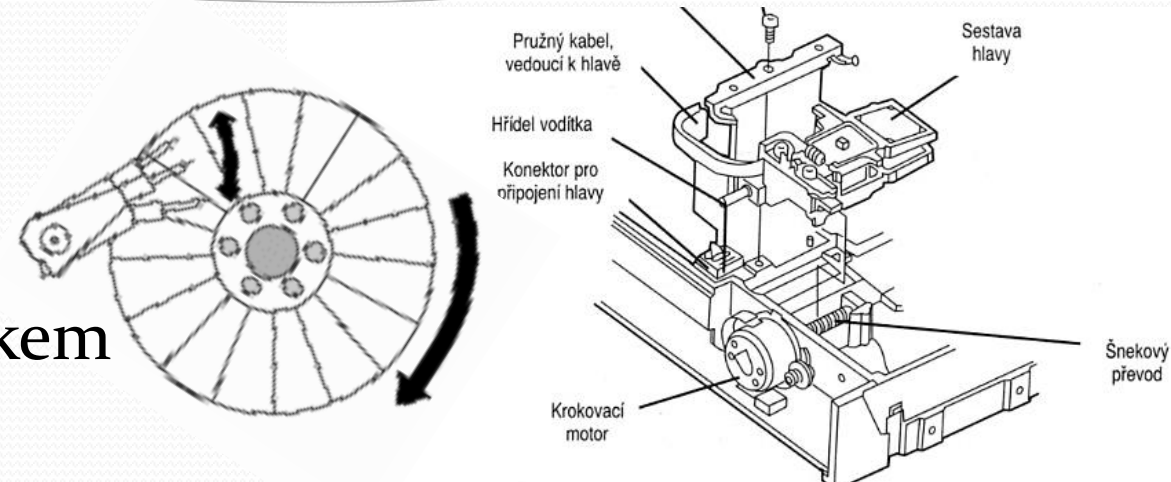
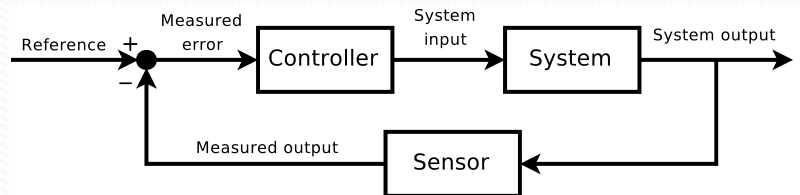
- MR hlava vychází ze změny odporu při průchodu magnetickým polem
 - Lze ji použít pouze pro čtení
 - Vysoká citlivost
 - Nutnost odstínění od okolních mag. polí
- Indukční TF hlava pro zápis
 - TF hlava zapisuje širší stopu než MR hlava čte
- PMR – kolmý zápis
- GMR hlavy – větší rozdíly při změně odporu
- TMR (tunneling magnetoresistance) - čtení

How head technology has enabled ever higher recording density



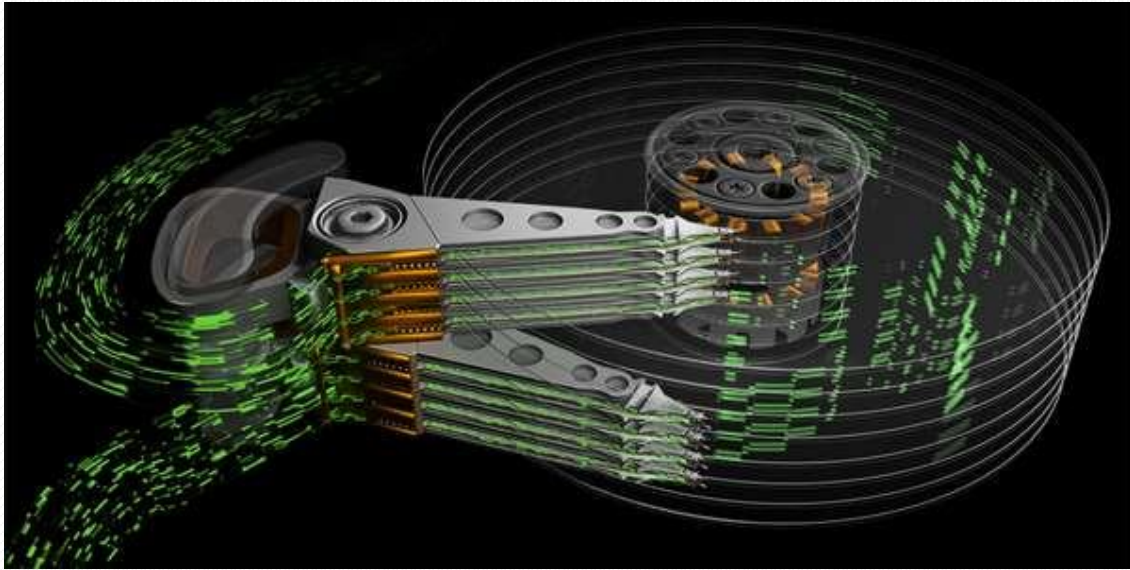
Pohony hlav

- Pohybuje závěsem s hlavami napříč diskem
- Elektromagnetický pohon
 - Vyšší přesnost, rychlost, spolehlivost
 - Dříve se používal krokový motor
- Pro přesné nastavení nad požadovanou stopu se používá zpětná vazba



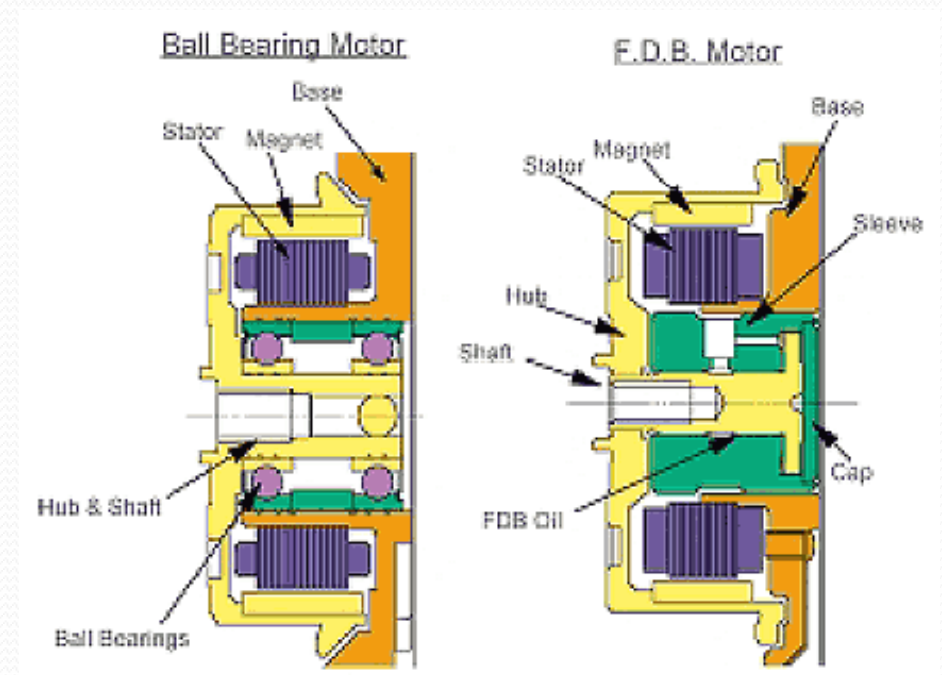
Hlavy pro čtení a záznam a rotační elektromagnetický pohon závěsu

Mach.2 (Seagate, 2021)



Pohon ploten disku

- Motorem přímo připojeným k hřídeli
 - Nevyužívá se převodů
 - Dříve vestavěn pod plotnami
 - Dnes vestavěny přímo v hřídeli disku
 - Bezvibrační chod
 - Fluidní ložiska
 - až 10x menší vůle, menší vibrace
- Přesné řízení rychlosti
 - Zajišťuje elektronika
 - Konstantní otáčky
 - Rychlost dána výrobcem – nejde měnit



Vzduchové filtry

- Disk je utěsněn (ne vzduchotěsně)
- Obsahuje dva typy filtrů
- **Recirkulační filtr**
 - Zachytávání malých částic
- **Barometrický (průchodový) filtr**
 - Vyrovnávání tlaků
 - Zachytávání částic a vlhkosti

Řídící deska (elektronika)

- polohování a pohon hlav
- pohon disku
- Předávání dat do řadiče
- Umístění Cache paměti - buffer
- u IDE disků integrován řadič
- u SCSI disků – řadič a adaptér SCSI sběrnice

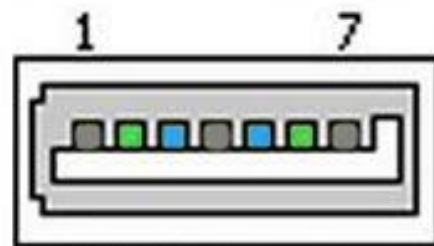
Kabely, konektory, propojovací prvky

- Napájecí konektor
 - 3.3 V, 5 V a 12 V
- Datový konektor
 - Záleží na typu disku
- Propojky (switch, jumper) pro nastavení disku

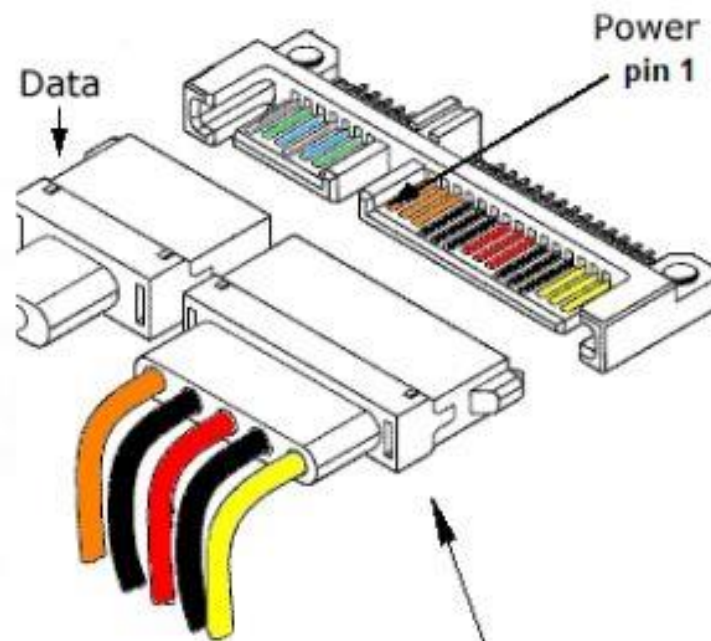
Konektory

SATA Data pinout

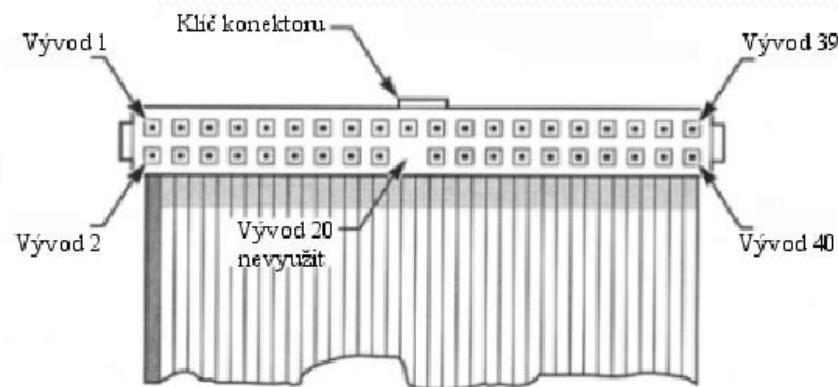
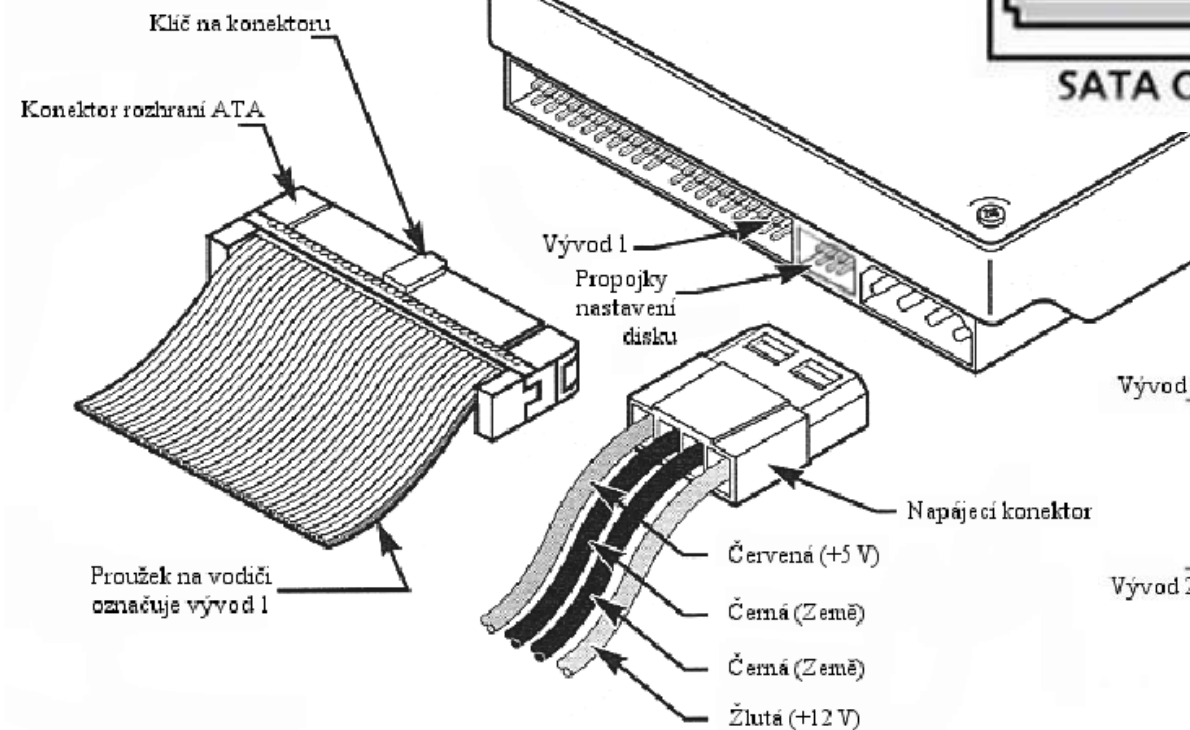
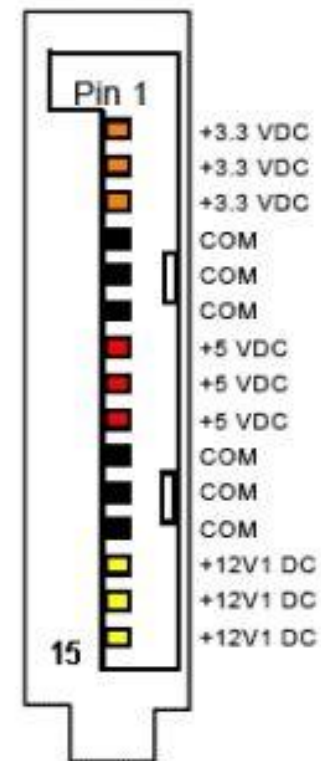
1	GND	Ground
2	A+	Transmit +
3	A-	Transmit -
4	GND	Ground
5	B-	Receive -
6	B+	Receive +
7	GND	Ground



SATA Connector



Serial ATA (SATA) Connector



Technologie S.M.A.R.T.

- Self-Monitoring, Analysis and Reporting Technology
- Předpovídání poruch pevných disků

Mezi sledované parametry patří:

Výška vznášení hlav nad povrchem ploten

Přenosová rychlost

Doba roztáčení ploten

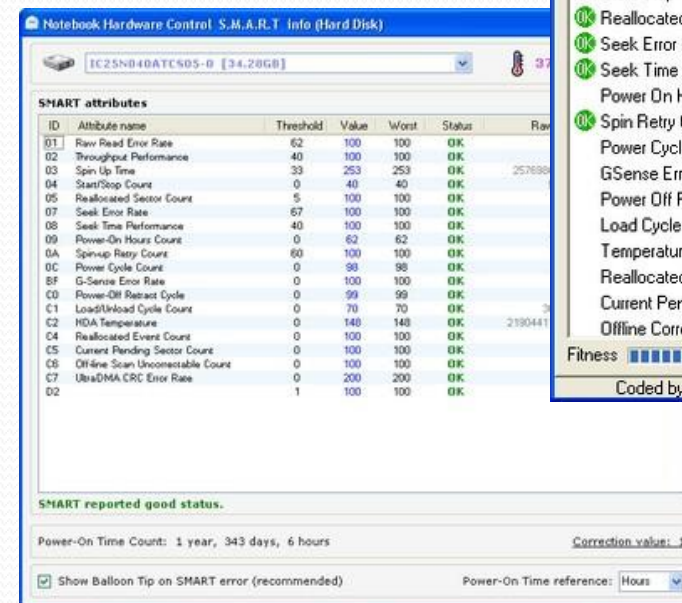
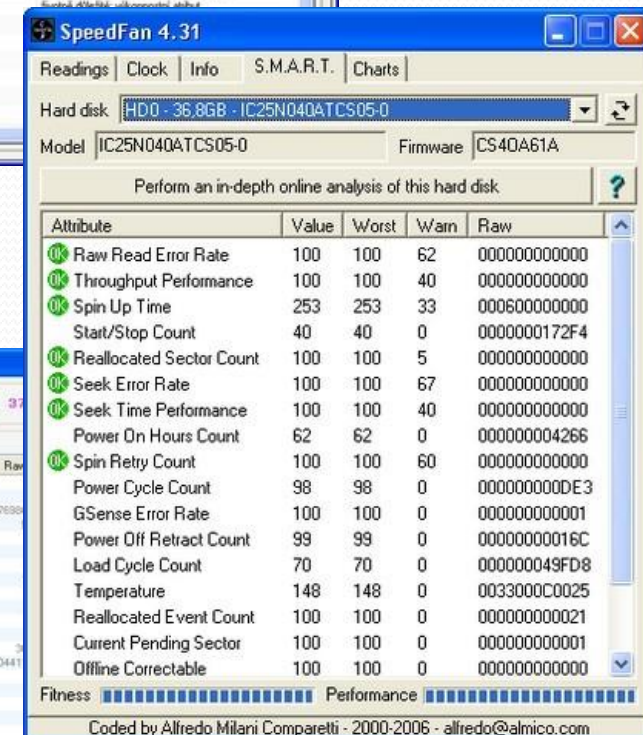
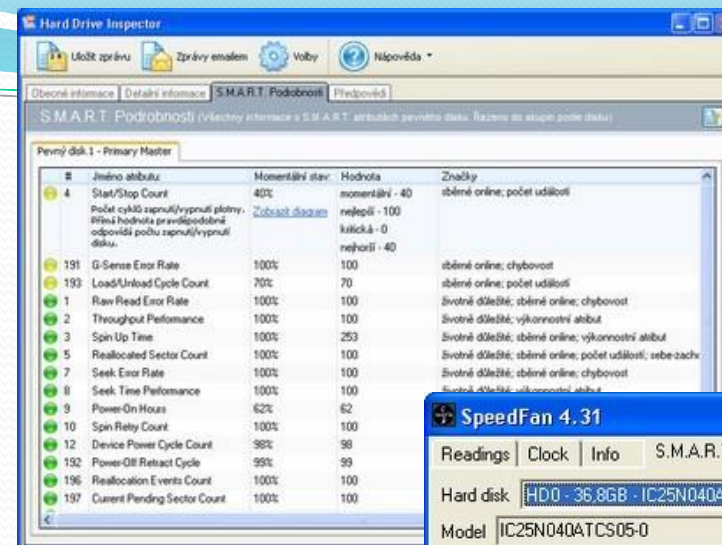
Počet realokovaných sektorů

Počet chyb při vyhledávání

Průměrná doba vyhledávání

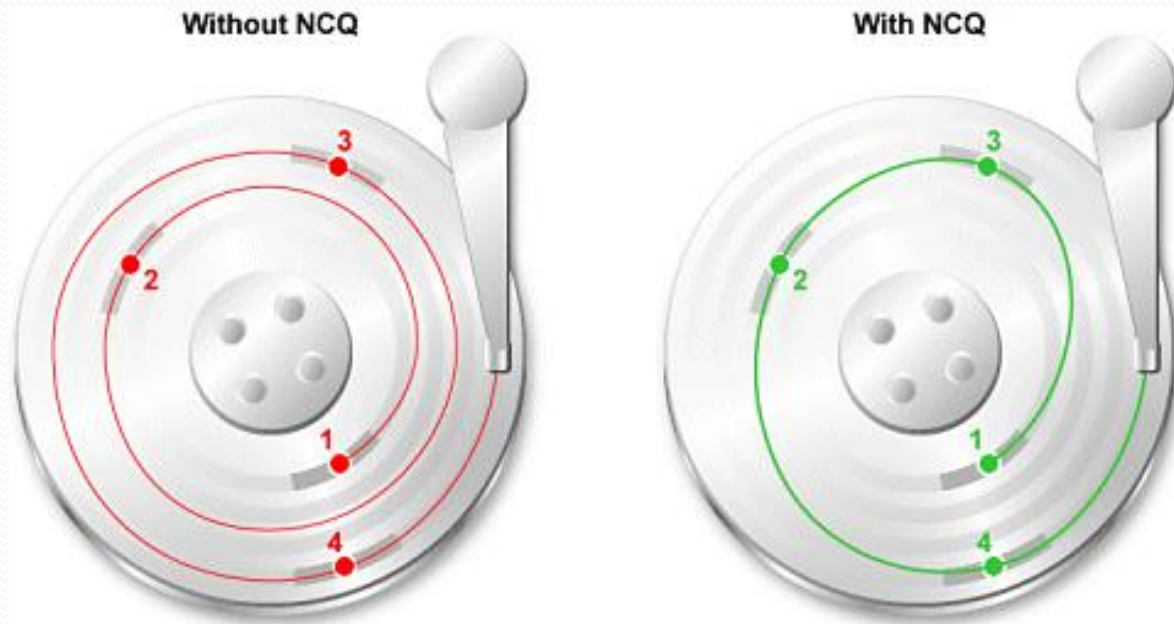
Počet opakovaných pokusů o roztočení ploten

Počet opakovaných pokusů o kalibraci disků



NCQ - Native Command Queuing

- Přirozené řazení příkazů
- HD optimalizuje pořadí požadavků
 - Redukovat nadbytečného pohybu hlaviček
 - Rychlejší + nižší opotřebení



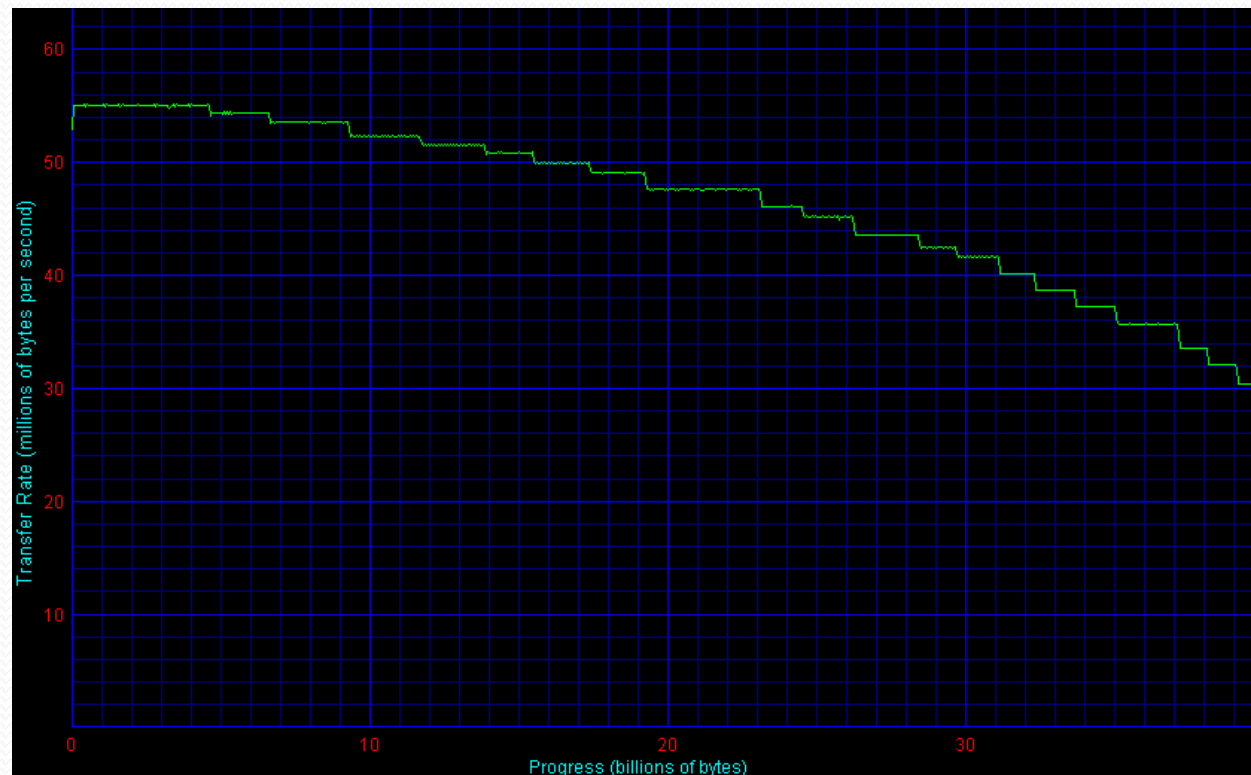
Without NCQ/TCQ	Head Moves
Movement from track 0 to track 8	8 tracks
Movement from track 8 to track 4	4 tracks
Movement from track 4 to track 7	3 tracks
Movement from track 7 to track 1	6 tracks
Total Distance of seeks	21 tracks
With NCQ/TCQ	Head Moves
Movement from track 0 to track 1	1 track
Movement from track 1 to track 4	3 tracks
Movement from track 4 to track 7	3 tracks
Movement from track 7 to track 8	1 track
Total Distance of seeks	8 tracks

Přístupová doba

- čas, který disku zabere vyhledat informaci
 - vystavování hlaviček
 - *track-to-track*
 - *full-stroke*
 - *average access*
 - rotační latence ploten

Sekvenční čtení

- Disk je rozdělen do několika zón s různým počtem sektorů na stopu
- Při konstantní rychlosti otáčení se liší počet sektorů na otáčku
- První sektor je na vnějším okraji



Parametry disku

• Kapacita	jednotky, desítky TB
• Velikost bufferu	desítky, stovky MB
• Rozhraní	SATA, SAS, ...
• Velikost sektoru	4 kB, 512 B (512e emulace navenek)
• Počet ploten, hlaviček	1-8
• Rychlost otáčení (RPM)	3 600-15 000
• Přístupová doba	4-12 ms
• Propustnost dat	> 100 MB/s
• Spotřeba (Idle & Read)	3-7 & 5-12 W
• MTBF	300 000 – 1 500 000 hodin
• Workload	55 -550 TB/rok
• Start/stop	50 000 – 300 000
• Operační teplota	5-60 °C
• Teplotní gradient	25 °C/hod.
• Operační/neoperační šok	60/300 g
• Záruka	2-5 let