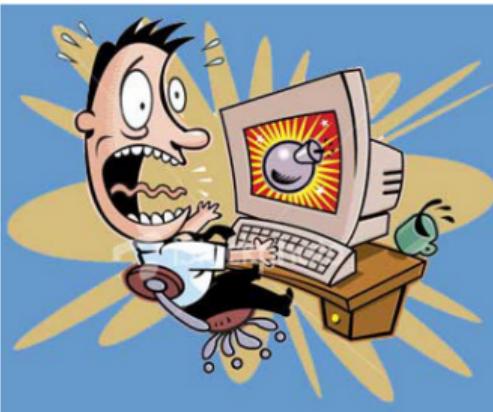
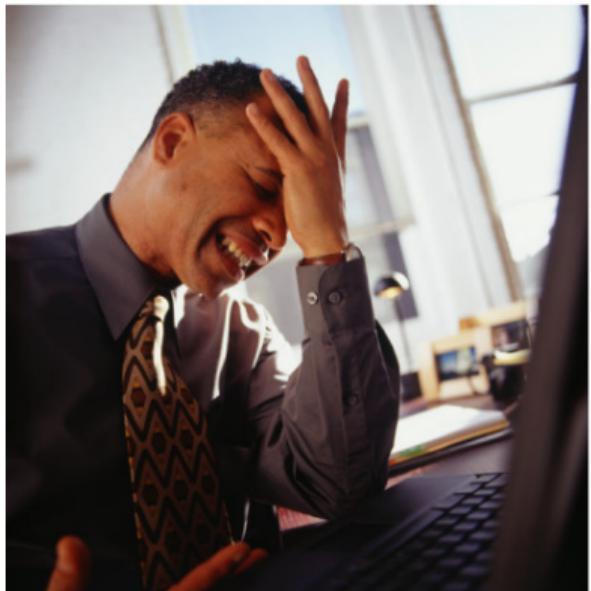


Zbliża się pora dobranocki



Drogie dzieci! Dziś opowiem Wam, jak za moich czasów przechowywaliśmy dane...

- To, co powiem, może być dla wielu osób przestarzałe, a przynajmniej staromodne.
- Wyrażone przeze mnie opinie będą subiektywne.
- Nie jestem typowym konsumentem technologii informacyjnej.



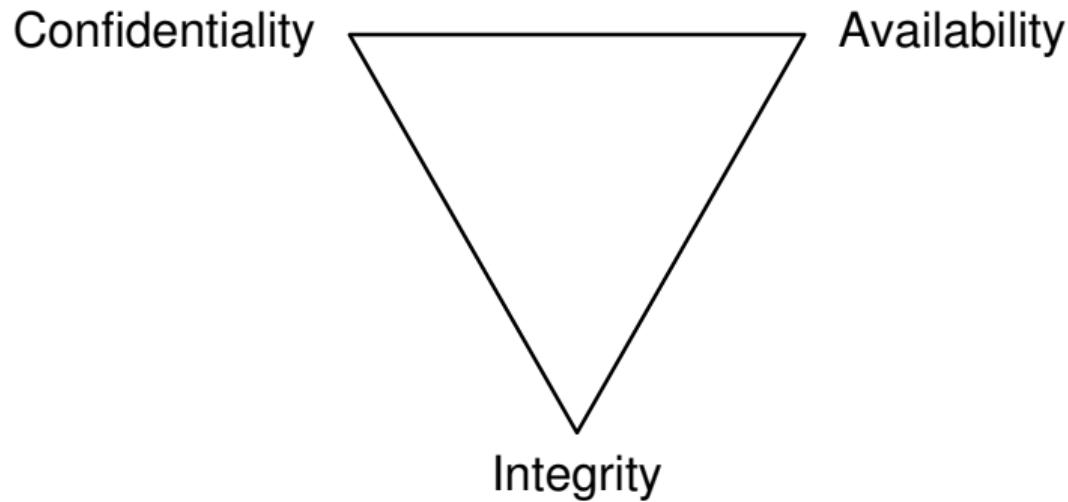
Kurs administrowania systemem Linux

Zajęcia nr 5a: Kopie zapasowe

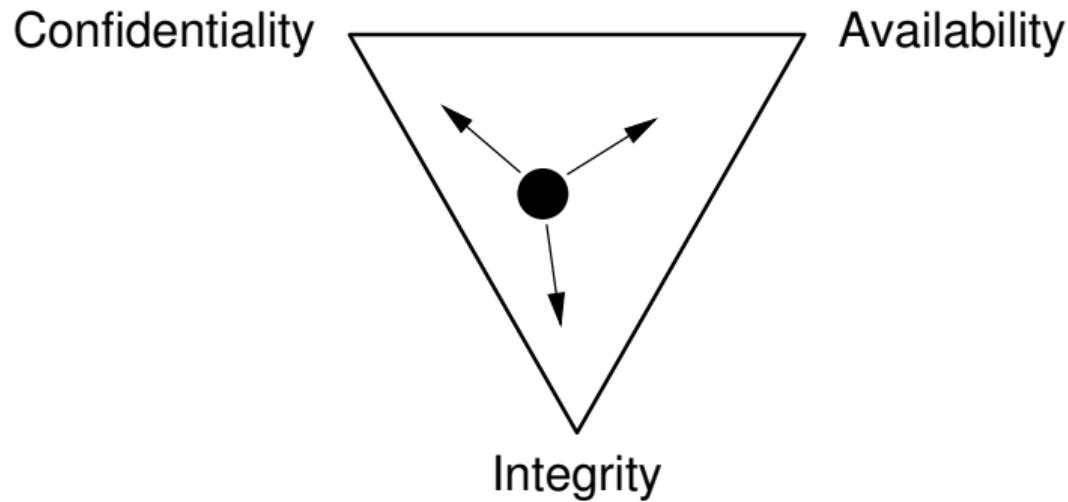
Instytut Informatyki Uniwersytetu Wrocławskiego

3 kwietnia 2024

Ochrona danych w spoczynku a triada CIA

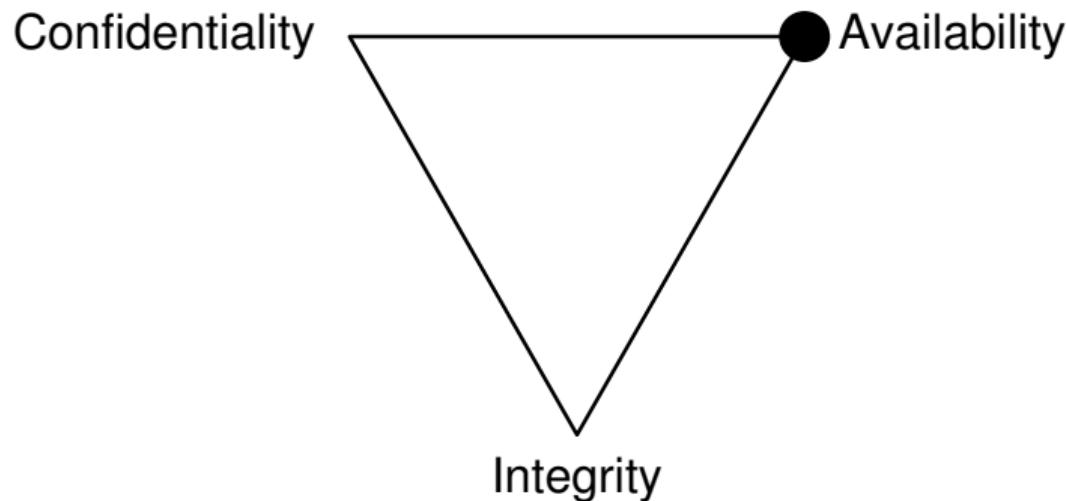


Ochrona danych w spoczynku a triada CIA



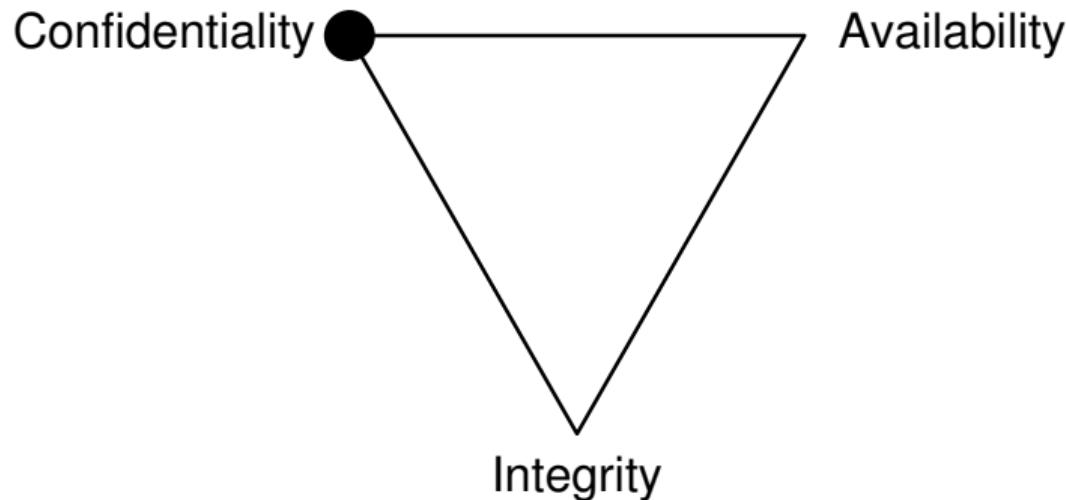
Ochrona danych w spoczynku a triada CIA

100% dostępności: zrób mnóstwo kopii
i roześlij je po świecie!



Ochrona danych w spoczynku a triada CIA

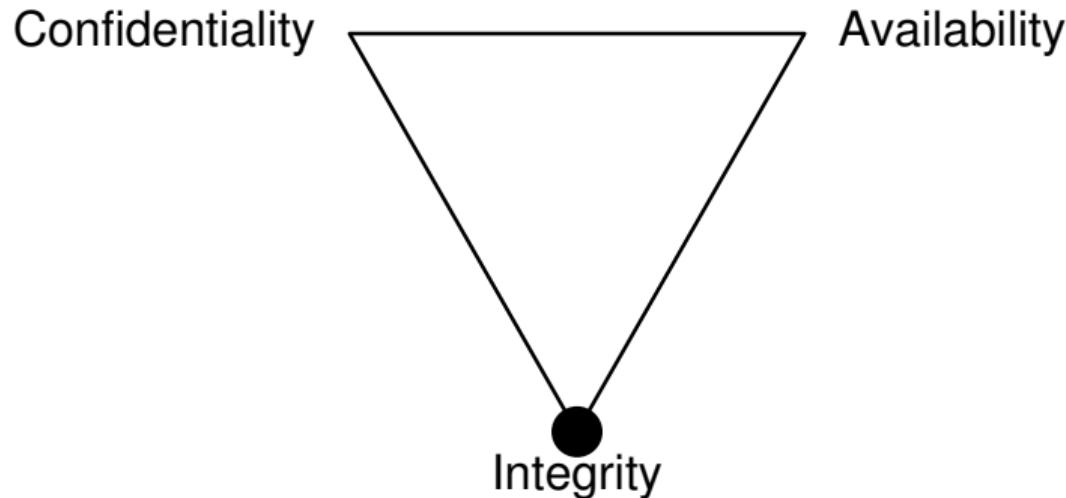
100% poufności: zniszcz wszystkie kopie!



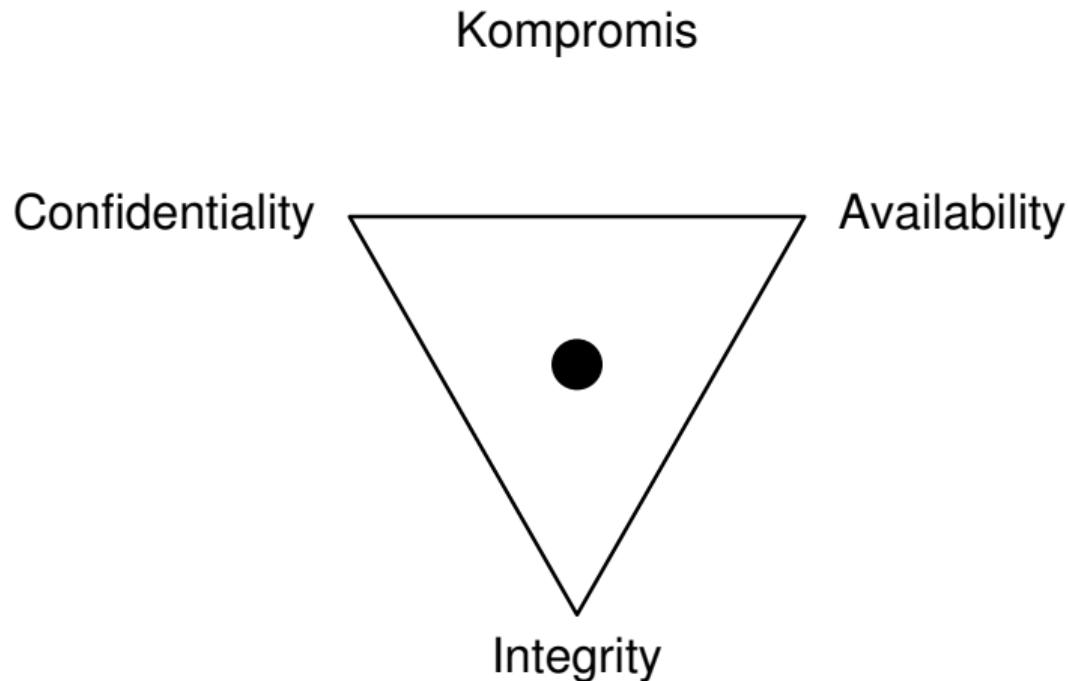
Ochrona danych w spoczynku a triada CIA

?

(Macie pomysły na głupie rozwiązanie?)



Ochrona danych w spoczynku a triada CIA



Shit happens.

(przysłowie staroangielskie)

Nie mogłem oddać zadania w terminie, bo zepsuł mi się komputer i straciłem cały projekt.

Arno Wagner o utracie danych

From observations on the cryptsetup mailing-list, people typically go through the usual stages of grief (Denial, Anger, Bargaining, Depression, Acceptance) when this happens to them. Observed times vary between 1 day and 2 weeks to complete the cycle. Seeking help on the mailing-list is fine. Even if we usually cannot help with getting back your data, most people found the feedback comforting.

Arno Wagner: Cryptsetup FAQ

DABDA model: Elisabeth Kübler-Ross, *On Death and Dying*, 1969.

Pamiętajcie: *shit happens!*



20 maja 2013

POKAZYWARKA

Błagam, błagam Wykopowe CSI o pomoc!

Jeden z chłopców pokazanych na zdjęciach w towarzystwie kolegi z tego samego zdjęcia ukradł mi dzisiaj laptopa z bezcennymi danymi!!!

Jestem w rozpaczy! Jesteście moją nadzieję, bo policja obiecała wziąć się do sprawy dopiero jutro.

W komputerze są bezcenne dane - rysunki w AutoCadzie, wizualizacje, umowy z klientami i zdjęcia z podróży.

W skradzionej razem z komputerem torbie były także dwa telefony, pendrivy, bardzo ważne notatniki, płyty CD.

Zawsze pilnowałem komputera jak żrenicy własnego oka, również i tym razem zabrałem go ze sobą z samochodu.

W Macu po zamówieniu kawy i kanapek usiadłem przy stoliku i czekałem na realizację zamówienia. Gdy kanapki były gotowe, pani z obsługi zwołała mnie, więc podszedłem do lady, a w tym momencie torba z zawartością zniknęła.

<http://pokazywarka.pl/j112hs/>

Pamiętajcie: *shit happens!*

Jego dom płonął, więc rzucił się w ogień, by uratować swoje dwie skończone książki

TS 18.09.2016 18:07



GAZETA.PL WIADOMOŚCI

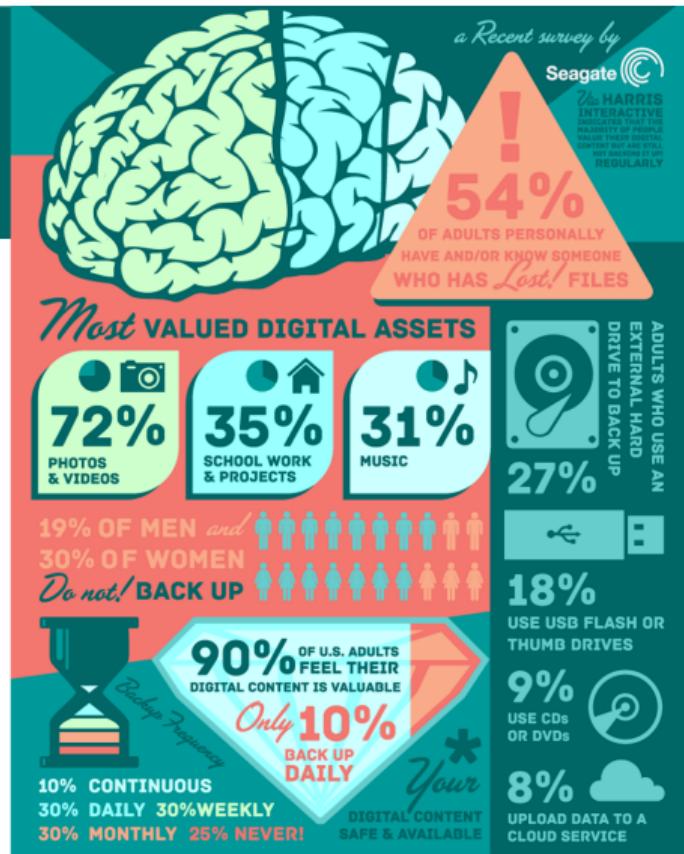


Gideon Hodge, pisarz z Nowego Orleanu, nie wahal się ani chwili, kiedy zobaczył, że jego dom stanął w ogniu. Miał tylko jeden cel - uratować laptop, w którym przechowywał dwie skończone książki.

- Tam w zasadzie był cały dorobek mojego życia - powiedział Gideon Hodge po tym, jak cały i zdrowy, z laptopem w rękach, opuścił swój płonący dom. 35-latek z Nowego Orleanu nie miał żadnych wątpliwości - kiedy zobaczył, że jego posiadłość stoi w ogniu, wiedział jedno: musiał uratować komputer.



Infografika Seagate (plus bonus)



Światowy Dzień Kopii Zapasowych (31 marca)



BACK IT UP

MARCH 31ST

Business files, photos, financial information—what if it all vanished? It's not a question of if your hard drive will fail, it's *when*. And if the data isn't backed up, it's gone.

World Backup Day is a day for people to learn about the role of data in our lives and the importance of regular backup.

REGULAR BACKUPS:

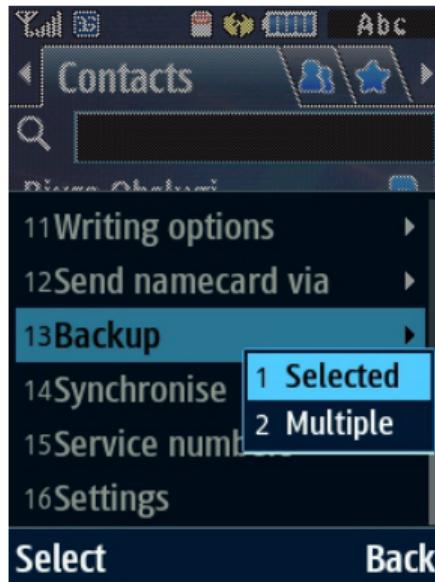
Why the effort? Turns out, not many people complete regular backups of their data.

In fact, less than 10% of people backup their data every day.

Frequency	Percentage
Daily	28%
Weekly	14%
Monthly	7%

<http://www.worldbackupday.com/>

Nawet „kontakte” w „dumbphonie” można skopiować



Książkę telefoniczną można przesyłać w formacie vCard (RFC 6350):

- na kartę SD telefonu (*backup*),
- na komputer poprzez łącze Bluetooth (*send via Bluetooth*),
- do adresata poczty elektronicznej (*send via e-mail*),
- do chmury (*synchronise* — uwaga na poufność!) używając otwartego standardu synchronizacji danych syncML (Open Mobile Alliance).

Uwaga: ten slajd jest celowo anachroniczny!

```
BEGIN:VCARD
VERSION:2.1
N;CHARSET=UTF-8;ENCODING=QUOTED-PRINTABLE:;Joe
FN;CHARSET=UTF-8;ENCODING=QUOTED-PRINTABLE:;Joe Biden
TEL;CELL; HOME; VOICE:0012024561111
END:VCARD
```

Teoria niezawodności

Mean time between failures (MTBF)

T — zmienna losowa opisująca moment awarii.

$$\text{MTBF} = ET.$$

Przeważnie przyjmuje się, że częstość awarii (*failure rate*) nie zależy od czasu (proces jest *bez pamięci*), tj. wynosi

$$\lambda = 1/\text{MTBF}.$$

Zmienna T ma wówczas rozkład *wykładniczy o dystrybuancie*

$$\Pr(T \leq t) = 1 - e^{-\lambda t}.$$

MTBF przeciętnych dysków mechanicznych wynosi 400 tysięcy godzin. Oznacza to, że prawdopodobieństwo, iż dysk ulegnie awarii w ciągu 5 lat użytkowania wynosi około 0.1.

Prawdopodobieństwo, że dysk ulegnie awarii w ciągu doby wynosi $6 \cdot 10^{-5}$, a w ciągu 2 tygodni — $8 \cdot 10^{-4}$.

Prawdopodobieństwo trafienia szóstki w Toto-Lotku wynosi $7.2 \cdot 10^{-6}$.

Ochrona danych przez powielanie

Rodzaje kopii

- Off-line (zwykle na nośnikach zewnętrznych)
- On-line (RAID, serwery backupowe, *hot standby*)

Rodzaje kopii off-line

- Archiwizacja
- Kopie zapasowe (*backupy*)

Zarządzanie kopiami

- pojedyncza aktualizowana kopia lustrzana
 - po synchronizacji niewłaściwe (np. omyłkowe) zmiany niemożliwe do anulowania
- wiele obrazów z różnych momentów czasu
 - wymaga dużej pojemności lub wielu nośników
 - można używać nośników z pojedynczym zapisem (np. dysków CD)
- pełna kontrola wersji
 - skomplikowany system, możliwy do zaimplementowania tylko dla niektórych rodzajów danych (np. pliki tekstowe)

Migawki (snapshots)

- LVM snapshot partitions
- ZFS snapshots

Systemy kontroli wersji

- RCS
- Subversion
- CVS
- Git

Disaster recovery

- Zbiór procedur, które w razie wystąpienia awarii pozwalają na utrzymanie/przywrócenie działalności.
- *Checklists*
- Automatyzacja
- Kopie zapasowe
 - procedury tworzenia kopii
 - procedury odtwarzania danych

Chaos monkey (Netflix @ AWS)



Koszt składowania danych

- płyty CD-R (700 MB; 1.40 zł/GB) i DVD±R (4.7 GB; 35 gr/GB)
 - przeważnie tylko jednokrotny zapis
 - ulegają procesom starzenia, bywają zawodne
- dyski mechaniczne: 0.5–22 TB (10–15 gr/GB)
 - nadal najlepsza opcja w zastosowaniach domowych
- dyski SSD: 0.12–4 TB (25 gr/GB)
- pendrive'y: 32–1000 GB (65 gr/GB)
- chmury: Dropbox Pro (25 gr/GB/rok)

Narzędzia

Narzędzia do kopiowania plików i katalogów

`cp` kopiuje pliki i katalogi (ważne opcje: `-a`, `-b`, `-i`, `-p`, `-r`).

`dd` „niskopoziomowe” kopiowanie plików.

Narzędzia do synchronizacji katalogów

`rsync` szybkie i wygodne narzędzie do kopiowania i synchronizacji.

`unison` niezawodne narzędzie do synchronizacji dwukierunkowej.

Narzędzia do nagrywania płyt CD/DVD

`wodim` program do nagrywania płyt

`genisoimage` program do tworzenia obrazów dysków

`growisofs` program do tworzenia i nagrywania obrazów dysków

`cdrdao` program do odczytu/zapisu na dyskach CD w trybie DAO

`readom` program do odczytu/zapisu na dyskach CD

`k3b` program do nagrywania płyt w KDE

Zinwentaryzuj swoje dane i opracuj system backupowania

1. Zidentyfikuj

- zawartość dysków komputerów stacjonarnych, laptopów;
- dane przechowywane na zewnętrznych serwerach, w chmurach itp.;
- urządzenia przenośne: telefony, tablety, czytniki e-booków, aparaty fotograficzne, itp.

2. Posegreguj

- Utrata których danych będzie nieodwracalna?
- Utrata których danych będzie kosztowna (czas, pieniądze)?
- Których danych nie warto backupować?

3. Uporządkuj

- Uwaga: błąd ludzki podczas porządkowania jest częstym powodem utraty danych!

„Jak trzymam w chmurze, to oni się zajmują backupem!”



r/DataHoarder • 1 yr. ago
New_Dragonfly9732



ALL my photos in mega.nz disappeared. Everything.

Question/Advice

I have uploaded photos in mega.nz in September 2019, when there was 50gb for the free plan, now I tried to access them, and they are ALL disappeared. There's literally nothing. I have read that that switched from 50gb to 15gb for free plan, so I expect to have some photos deleted but the remaining 15gb still there, but again, there is nothing. Why? Can I do something?



tyroswork • 1y ago

Just restore from backup and don't use mega.nz anymore. You do have a backup, right?



↑ 21 ↓

Share ...

Plan backupowania

Katalog	Ścieżka	[MB]	Backup
Pliki „robocze” (w tym bieżąca poczta)	do wyboru	1 122	raz na godzinę
Pliki konfiguracyjne systemu	/etc	12	rzadziej, ew. po zm.
Pliki konfiguracyjne użytkownika	fragm. /home/user	2	rzadziej, ew. po zm.
Własne zdjęcia, dawne „robocze” itp.	do wyboru	26 902	rzadziej, ew. po zm.
Artykuły, książki, zdjęcia, filmy z Internetu	do wyboru	128 219	rzadziej, ew. po zm.
Bkp drobnego sprzętu (dane i firmware)	do wyboru	37 421	rzadziej, ew. po zm.
Katalog domowy	reszta /home/user	4 321	nigdy
System, część stała	/usr	9 201	nigdy
System, część zmienna	/var	3 285	nigdy

Backup plików „roboczych” trwa 22 s, backup całości — 3 m 30 s.

Unison — synchronizacja dwukierunkowa



- Napisany w OCamlu przez Bejamina Pierce'a, Trevora Jima i Jérôme Vouillona w 2001 roku.
- Benjamin C. Pierce, Jérôme Vouillon, *What's in Unison? A Formal Specification and Reference Implementation of a File Synchronizer*, TR MS-CIS-03-36, UPenn, 2004.
- Formalna specyfikacja protokołu i dowód poprawności implementacji w Coq.
- Motto Hipokratesa: „Po pierwsze nie szkodzić”.

Plik konfiguracyjny /home/user/.unison/robocze.prf

```
auto=true
times=true
logfile=/tmp/robocze_unison.log
root=/home/user/robocze
root=/media/bkpdsk/robocze
```

Uruchamianie:

```
linux> unison robocze
```

Dysk backupowy powinien być zamontowany w katalogu /media/bkpdsk i powinien być w nim założony katalog /robocze.

Montowanie urządzeń

```
linux# mkfs.ext4 -L BKP_PRIV /dev/sdb
```

```
linux# blkid /dev/sdb
```

```
UUID="716a1218-5f1b-4851-b0cb-be3833dd22e1" TYPE="ext4"
```

```
linux# findfs UUID="716a1218-5f1b-4851-b0cb-be3833dd22e1"  
/dev/sdb
```

```
linux# findfs LABEL=BKP_PRIV  
/dev/sdb
```

Plik /etc/fstab

```
UUID=716a1218-5f1b-4851-b0cb-be3833dd22e1 /media/bkp_priv ext4 noauto 0 0  
LABEL=BKP_PRIV /media/bkp_priv ext4 noauto 0 0
```

Skrypty i automatyzacja gwarancją niezawodności

Skrypt mkbkp

```
findfs LABEL=BKP_PRIV || exit 1
mount /media/bkp_priv
unison priv
umount /media/bkp_priv
```

Synchronizacja jednokierunkowa: rsync

Synchronizowanie czytnika ebooków

```
/usr/bin/rsync \  
    --verbose --human-readable --progress \  
    --recursive \  
    --times --omit-dir-times --links --safe-links \  
    --no-owner --no-perms --no-group \  
    --modify-window=2 --delete \  
    /home/user/ebooks/ /media/kindle/
```

Stan zdrowia dysku

Dawniej

Polecenie `badblocks` sprawdzało, które sektory są uszkodzone i generowało listę bloków (ciągów sektorów), które system plików miał omijać. Ze względu na realokację sektorów wewnętrz dysków, używanie `badblocks` jest obecnie bezcelowe.

Obecnie

Standard ATA w wersji 3 (ATA-3, inaczej EIDE, 1997) wprowadził *feature set S.M.A.R.T. — Self-Monitoring, Analysis and Reporting Technology*:

- m.in. komendy SMART DISABLE/ENABLE OPERATIONS, SMART EXECUTE OFF-LINE IMMEDIATE, SMART READ DATA/LOG
- standard nie określa precyzyjnie znaczenia atrybutów S.M.A.R.T. (są „*vendor specific*”)

S.M.A.R.T.: polecenie smartctl

ID#	ATTRIBUTE_NAME	FLAGS	VALUE	WORST	THRESH	FAIL	RAW_VALUE
1	Raw_Read_Error_Rate	POSR--	099	099	051	-	108
3	Spin_Up_Time	POS---	094	094	011	-	2830
4	Start_Stop_Count	-O--CK	098	098	000	-	1691
5	Reallocated_Sector_Ct	PO--CK	100	100	010	-	0
7	Seek_Error_Rate	POSR--	100	100	051	-	0
8	Seek_Time_Performance	P-S--K	100	100	015	-	0
9	Power_On_Hours	-O--CK	093	093	000	-	35400
10	Spin_Retry_Count	PO--CK	100	100	051	-	0
11	Calibration_Retry_Count	-O--C-	100	100	000	-	0
12	Power_Cycle_Count	-O--CK	098	098	000	-	1691
13	Read_Soft_Error_Rate	-OSR--	099	099	000	-	108
183	Runtime_Bad_Block	-O--CK	100	100	000	-	0
184	End-to-End_Error	PO--CK	100	100	099	-	0
187	Reported_Uncorrect	-O--CK	100	100	000	-	308
188	Command_Timeout	-O--CK	100	100	000	-	0
190	Airflow_Temperature_Cel	-O--K	079	069	000	-	21 (Min/Max 13/21)
194	Temperature_Celsius	-O---K	074	067	000	-	26 (Min/Max 13/26)
195	Hardware_ECC_Recovered	-O-RC-	100	100	000	-	2448
196	Reallocated_Event_Count	-O--CK	100	100	000	-	0
197	Current_Pending_Sector	-O--C-	100	100	000	-	12
198	Offline_Uncorrectable	----CK	100	100	000	-	0
199	UDMA_CRC_Error_Count	-OSRCK	100	100	000	-	0
200	Multi_Zone_Error_Rate	-O-R--	100	100	000	-	0
201	Soft_Read_Error_Rate	-O-R--	253	253	000	-	0
		_	K	auto-keep			
		_	C	event count			
		_	R	error rate			
		__	S	speed/performance			
		__	O	updated online			
		__	P	prefailure warning			

S.M.A.R.T. error log

Error 122 [1] occurred at disk power-on lifetime: 31074 hours (1294 days + 18 hours)
When the command that caused the error occurred, the device was active or idle.

After command completion occurred, registers were:

ER -- ST COUNT LBA_48 LH LM LL DV DC

-- - - - == - == == - - - - - - - -

40 -- 51 00 00 00 00 00 07 93 9a e0 00 Error: UNC at LBA = 0x0007939a = 496538

Commands leading to the command that caused the error were:

CR FEATR	COUNT	LBA_48	LH	LM	LL	DV	DC	Powered_Up_Time	Command/Feature_Name
----------	-------	--------	----	----	----	----	----	-----------------	----------------------

--	--	--	--	--	--	--	--	-----	-----
----	----	----	----	----	----	----	----	-------	-------

c8 00 00 00 08 00 00 00 07 93 98 e0 08 23:39:11.620 READ DMA

27 00 00 00 00 00 00 00 00 00 00 e0 08 23:39:11.620 READ NATIVE MAX ADDRESS EXT [OBS-ACS-3]

ec 00 00 00 00 00 00 00 00 00 00 a0 08 23:39:11.610 IDENTIFY DEVICE

ef 00 03 00 46 00 00 00 00 00 00 a0 08 23:39:11.610 SET FEATURES [Set transfer mode]

27 00 00 00 00 00 00 00 00 00 e0 08 23:39:11.610 READ NATIVE MAX ADDRESS EXT [OBS-ACS-3]

Device	Boot	Start	End	Sectors	Size	Id	Type
/dev/sda1	*	63	498014	497952	243.1M	83	Linux /boot
/dev/sda2		498015	625137344	624639330	297.9G	83	Linux LUKS /

S.M.A.R.T.: demon smartd

To: root@localhost
Subject: SMART error (CurrentPendingSector) detected on host
From: root <root@localhost>
Date: Mon, 01 Dec 2014 17:54:20 +0100

This message was generated by the smartd daemon running on:

host name: localhost
DNS domain: localdomain

The following warning/error was logged by the smartd daemon:

Device: /dev/sda [SAT], 9 Currently unreadable (pending) sectors

Device info: SAMSUNG HD322HJ, 320 GB

For details see host's SYSLOG.

You can also use the smartctl utility for further investigation.
The original message about this issue was sent at Thu Nov 13
19:42:16 2014 CET.

Another message will be sent in 24 hours if the problem persists.

Uszkodzone sektory

```
$ dd if=/dev/sda of=/dev/null count=1 skip=496537
dd: error reading '/dev/sda': Input/output error
0+0 records in
0+0 records out
0 bytes (0 B) copied, 1.94465 s, 0.0 kB/s
```

```
$ hdparm --read-sector 496538 /dev/sda
```

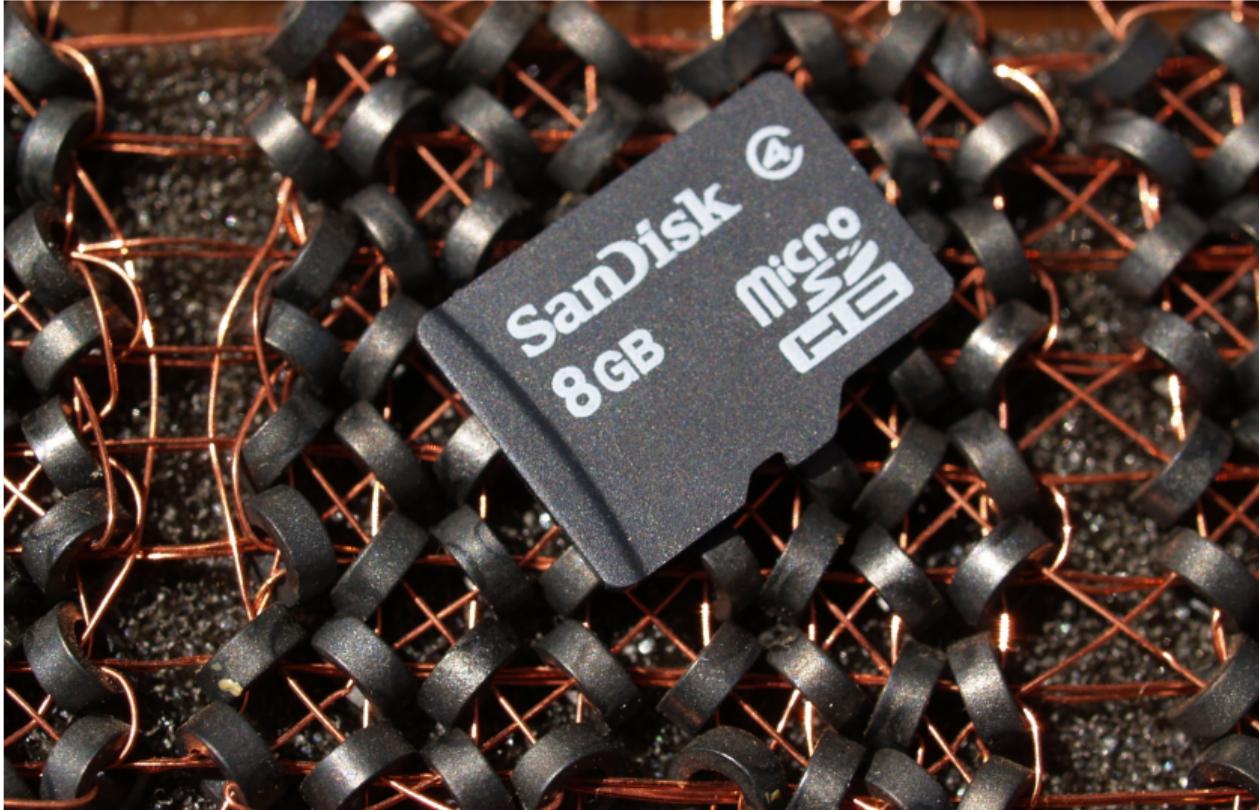
```
/dev/sda:
reading sector 496538: FAILED: Input/output error
```

Rozwiązanie:

- Overprovisioning.
- Mapa sektorów i ich realokacja.

Uwaga: Sektory zmienione przez realokację stają się niedostępne i nie można ich zamazać!!!

Pamięci komputerów



8B vs. 8GB

© Wikimedia Commons

Pamięci komputerów



Pamięć	
operacyjna	masowa
ulotna	nieulotna
szybsza	wolniejsza
mniej pojemna	bardziej pojemna
połączona bezpośrednio z procesorem	urządzenie peryferyjne
dostęp bezpośredni	urządzenie blokowe lub znakowe

Współczesne serwerownie



Best servers 2021

Server	FF	Sockets	Max Processors	Max Memory	Drives Supported	Starting Price
Cisco C225 M6	2U	2	3rd Gen Intel Xeon	12 TB	28	–
Dell EMC R750	2U	2	3rd Gen Intel Xeon	8 TB	24	–
Fujitsu RX4770	2U	4	3rd Gen Intel Xeon	15 TB	24	\$2,102
HPE DL380 Gen10	2U	2	2nd Gen Intel Xeon	9 TB	30	\$1,580
Huawei 2288H V5	2U	2	2nd Gen Intel Xeon	3 TB	31	\$1,900
IBM S922	2U	2	POWER9	4 TB	14	\$10,512
Inspur NF8480 M6	4U	4	3rd Gen Intel Xeon	12 TB	50	–
Lenovo SR630	1U	2	3rd Gen Intel Xeon	6 TB	16	\$2,285
Oracle X8-2	1U	2	2nd Gen Intel Xeon	64 TB	8	\$8,800
Dell EMC R840	2U	4	2nd Gen Intel Xeon	15 TB	26	\$7,706
HPE DL325 Gen10	1U	1	2nd Gen AMD EPYC	2 TB	8	\$1,639
IBM E980	4U	4	POWER9	64 TB	16	–
Oracle T8-1	1U	1	SPARC M8	1 TB	8	–



Trwałe nośniki danych

Mechaniczny †

- karty perforowane †
- taśmy perforowane †

Optyczny

- CD-ROM
- CD-R, CD-RW
- DVD-ROM
- DVD±R, DVD±RW
- DVD-RAM
- Blue-ray Disc

Magnetyczny

- pamięci bębnowe †
- pamięci rdzeniowe †
- dyskietki †
- taśmy magnetyczne
- dyski twarde

Półprzewodnikowy

- PROM, FPROM
- OTP NVM, NVRAM
- EPROM, EEPROM
- NOR flash, NAND flash

Obecnie dane są przeważnie składowane na *dyskach*

Urządzenia peryferyjne komputerów

- Urządzenia szeregowe — dane przesyłane w pojedynczych bajtach, dostęp szeregowy.
- Urządzenia blokowe — dane przesyłane w blokach ustalonego rozmiaru, dostęp bezpośredni (adresowanie).

Dyski twardy

- Dyski to urządzenia blokowe.
- Bloki są na dyskach zapisywane w *sektorach*.
- Rozmiar sektora: tradycyjnie 512 bajtów, w nowych dyskach 4096 bajtów (*advanced format*).
- Dyski mechaniczne, *Solid State Drive* (SSD) i hybrydowe.

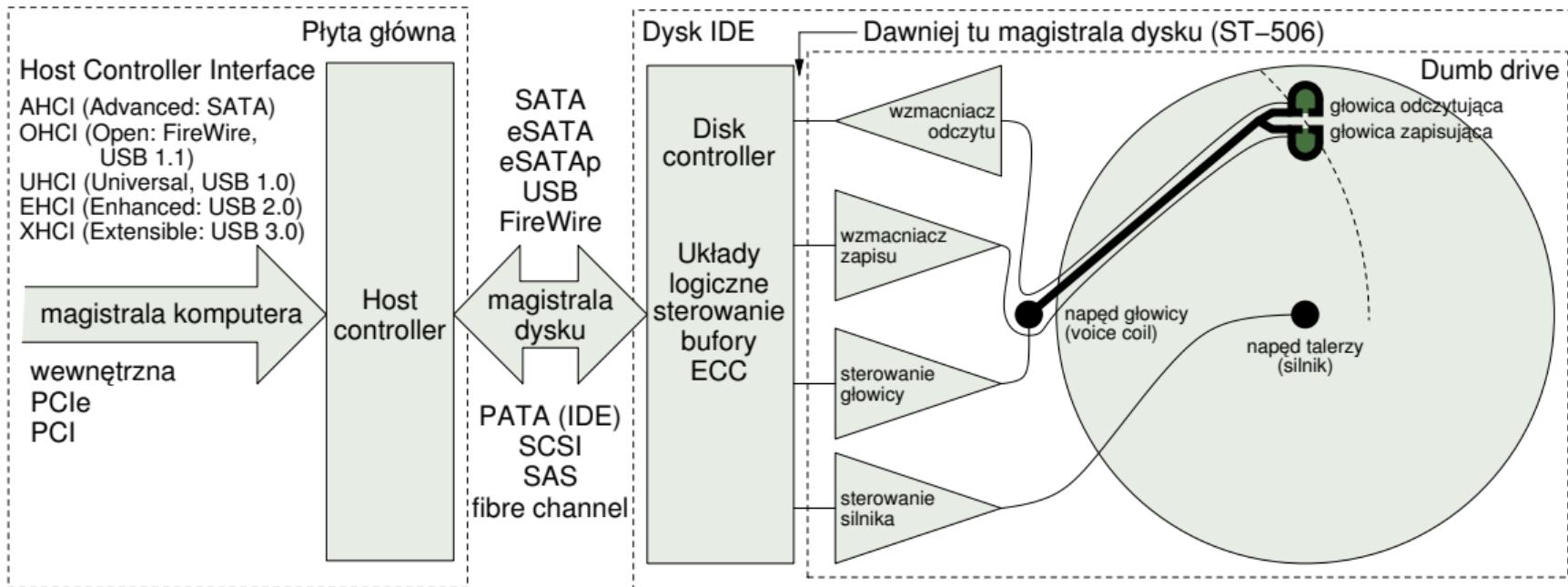
Dyski twardy mechaniczne (magnetyczne)

- Zapis sektora: *nagłówek, obszar danych* (512B lub 4096B), *suma kontrolna* (ECC, rzędu 50B).
- Podobnie jak w CD i DVD bardzo silna korekcja błędów za pomocą kodów Reeda–Solomona, LDPC itp.

Dyski twarde magnetyczne



Dyski twarde magnetyczne



Dumb disk

- *Host controller* zawiera kontroler dysku
- Bardzo prosta elektronika dysku
- Skomplikowana magistrala dysku

IDE disk (Integrated Drive Electronics)

- Kontroler dysku zintegrowany z elektroniką dysku
- Rozbudowana elektronika dysku
- Wyskopiowa magistrala dysku

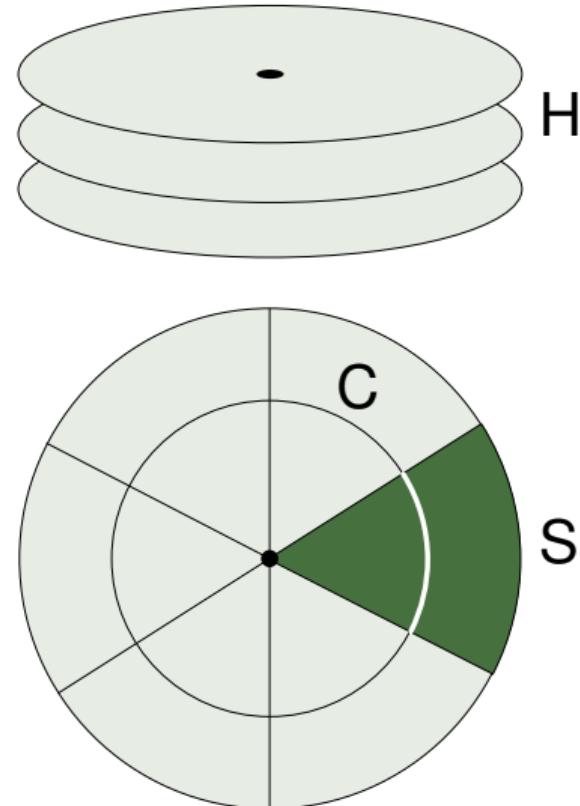
Numeracja sektorów klasycznego dysku magnetycznego

- Low-level formatting — struktura sektorów zapisana przez servowriter przed zmontowaniem dysku.
- Adres sektora: C/H/S (*cylinder, głowica, sektor*).
- Logical block addressing — liniowa numeracja sektorów.
- Ważne: Ze względu na duży *seek time* sąsiednie sektory powinny mieć kolejne adresy LBA. Dlatego w większości przypadków sektor zajmuje stałe miejsce na powierzchni plateru.

Logical block addressing (LBA)

$$\text{LBA}(C, H, S) = (C \cdot N_h + H) \cdot N_s + (S - 1)$$

Obecnie dyski mają zmienną liczbę sektorów na ścieżkę i numeracja CHS jest niedostępna.



Gęstość zapisu

Constant Angular Velocity (CAV)

- Stosowane w dyskach magnetycznych, dawniej w gramofonach (78, 45, $33\frac{1}{3}$ obr/min.)
- W przypadku stałego *bit rate'u* daje zmienną gęstość zapisu.

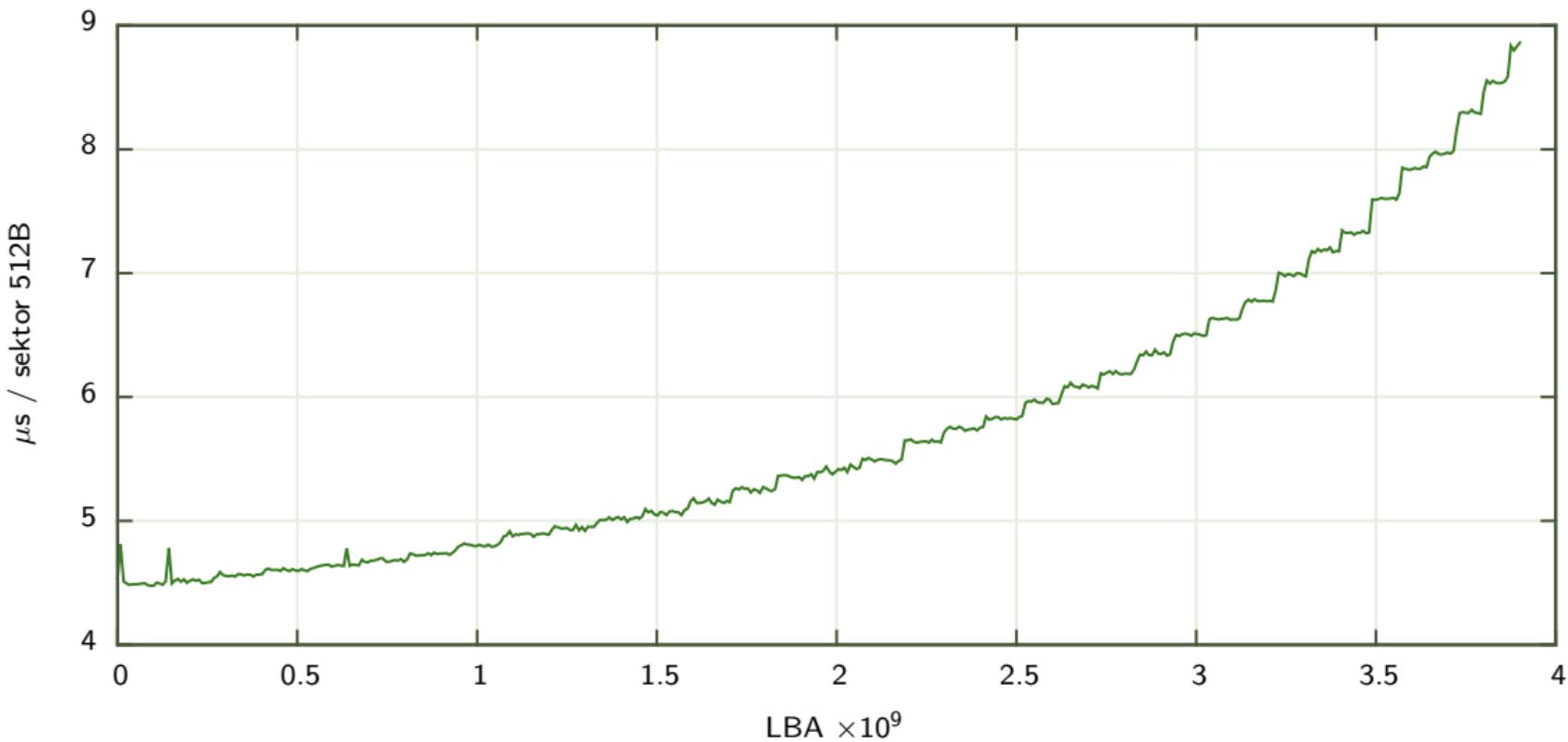
Constant Linear Velocity (CLV)

- Chichester Bell, Charles Tainter, *Recording and reproducing speech and other sounds*, US Patent No. 341,214, 1886.
- Stosowane obecnie w dyskach optycznych.
- Wymaga zmiany prędkości obrotowej w zależności od położenia głowicy.
- Ze względu na bezwładność platerów wydłuża *seek time*.

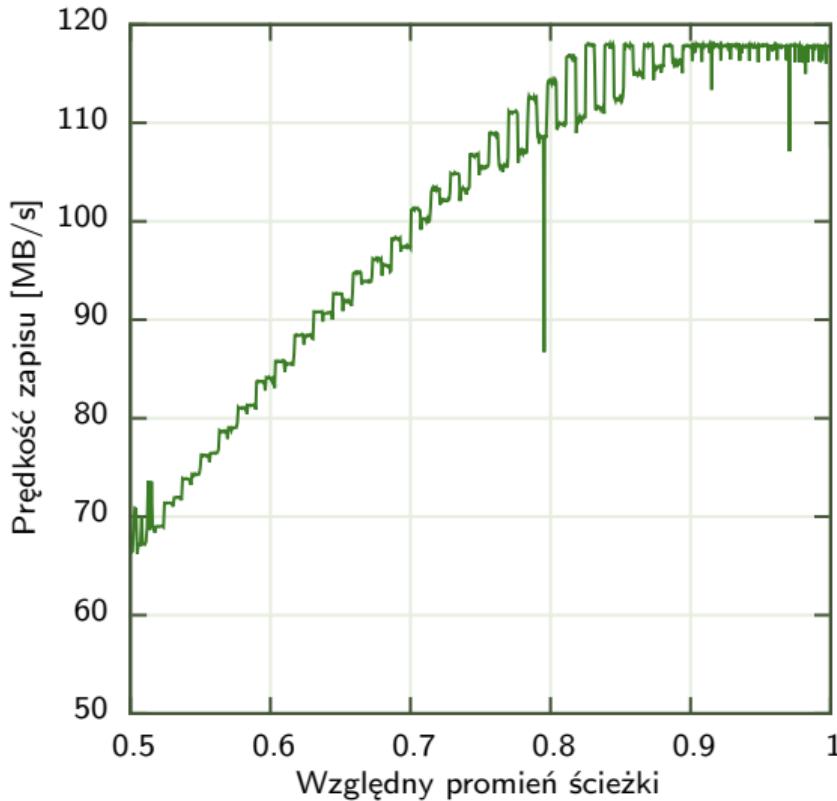
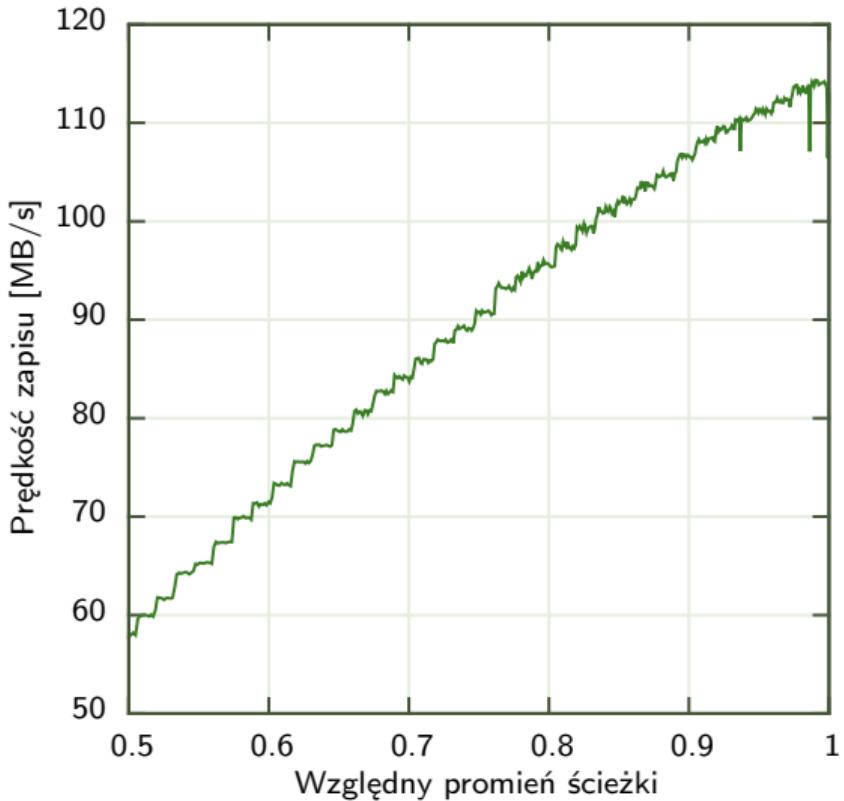
Zoned Bit Recording (ZBR)

- Charles I. Peddle, US Patent, 1962.
- Zmienna liczba sektorów na ścieżce przy zapisie CAV.
- Mniej więcej stała gęstość zapisu.
- Zmienny *bit rate* (na zewnątrz plateru — większy).

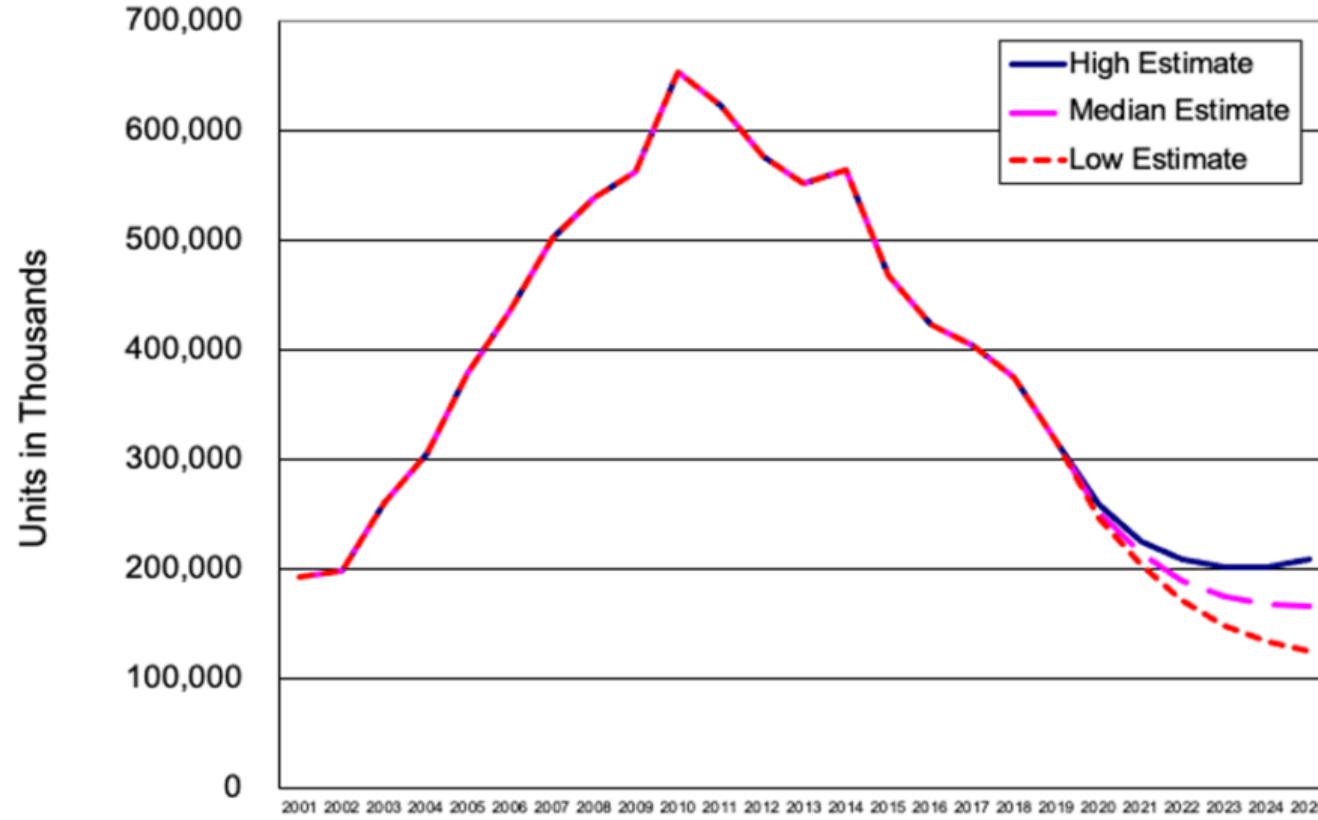
Czas zapisu sektora na dysku z ZBR (Toshiba Canvio Basic 2TB)



Prędkość zapisu na dysku CMR i SMR



Produkcja dysków magnetycznych (obecnie około 1 ZB rocznie)



Magistrale dysków

sprzęt	'80	'90	obecnie
konsumentcki	ST-506 (1980)	Parallel ATA (1994)	SATA (2003) USB (1995)
profesjonalny	SASI (1981)	(parallel) SCSI (1986)	SAS (2004) SAN

Magistrale dysków

IBM PC (*IBM Personal Computer*)

- 1. gen.: PC/XT (*X-tended Tech.*, 1983)
- 2. gen.: PC/AT (*Advanced Tech.*, 1984)

ATA (*AT Attachement*)

- początkowo tylko magistrala równoległa (obecnie PATA, dawniej „IDE”)
- obecnie **SATA** (*Serial ATA*)
- tylko dyski, inne urządzenia, np. CD ROM: **ATAPI** (*ATA Packet Interface*)
- wewnętrzna; zewnętrzne: **eSATA**, **eSATAp** (*Powered External SATA*)
- Nowe: **mSATA**, **M.2 SATA**

NVMe (Non-Volatile Memory express)

SCSI (*Small Computer System Interconnect*)

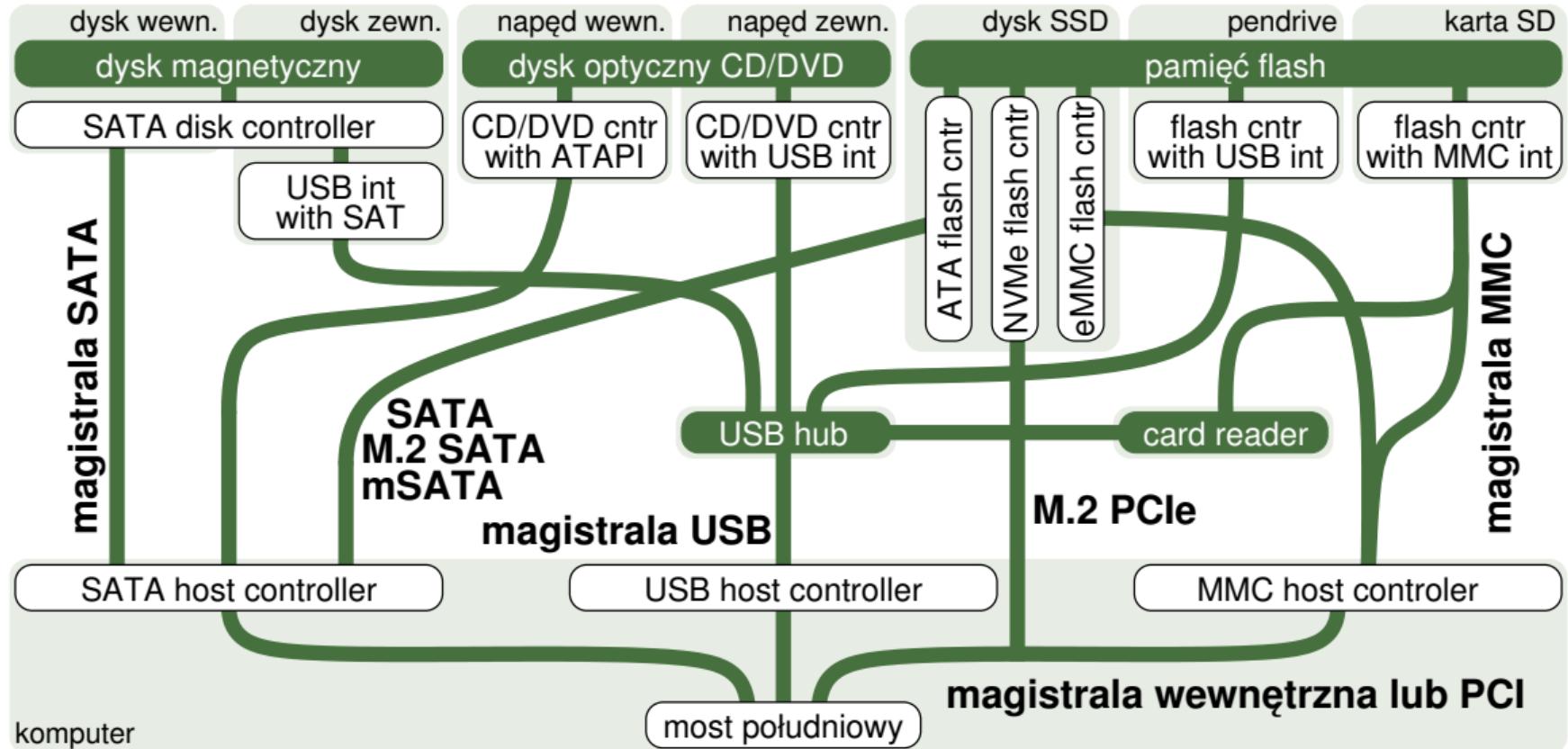
- początkowo tylko magistrala równoległa
- obecnie **SAS** (*Serial Attached SCSI*)
- wiele typów urządzeń

USB (*Universal Serial Bus*)

- zewnętrzna
- wiele typów urządzeń; klasy urządzeń dyskowych:
CBI (*Control/Bulk/Interrupt*), BBB (*Bulk/Bulk/Bulk*), **BOT** (*Bulk Only Transport*), **UAS** (*USB Attached SCSI*)

Inne: **FireWire**, **FibreChannel**, **InfiniBand**, **SAN** (*Storage Area Network*)

Urządzenia dyskowe: hardware



Protokoły magistral dysków konsumenckich: ATA

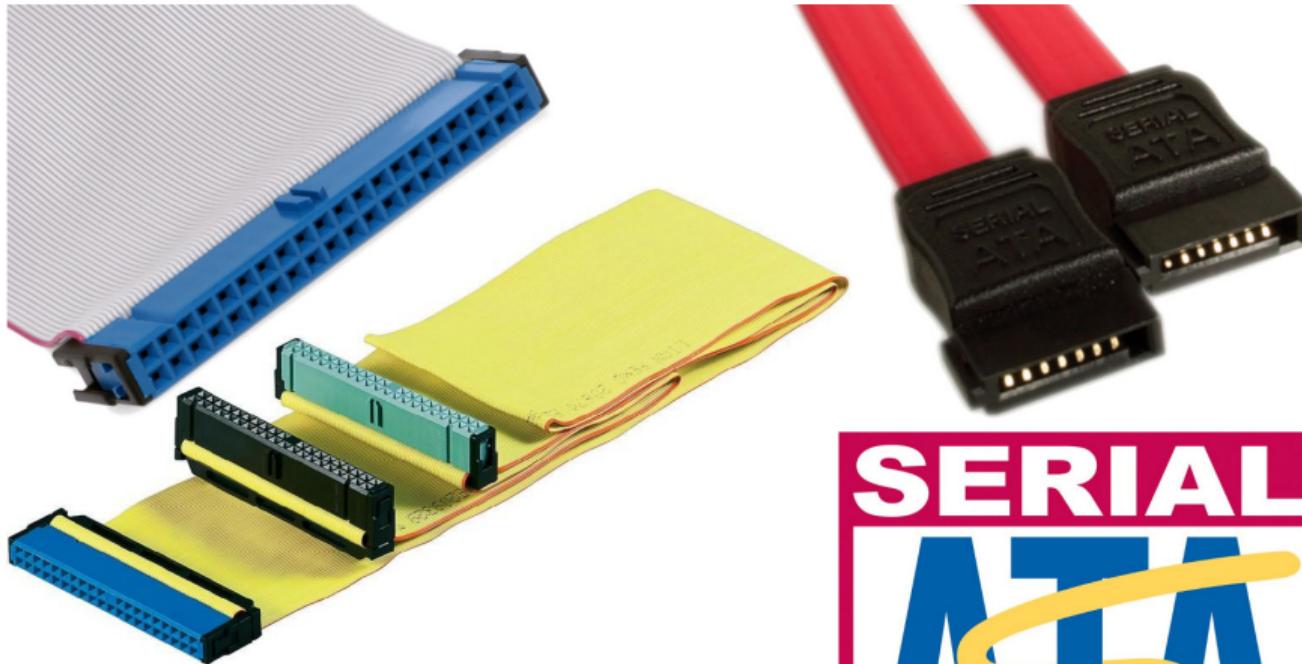
Standardy AT Attachment

Wersja	Standard ANSI	Wybrane możliwości
1	X3.221-1994*	
2	X3.279-1996*	Identify Drive command
3	X3.298-1997*	S.M.A.R.T., Security Command Set
4	NCITS 317-1998	ATAPI, Host Protected Area (HPA)
5	NCITS 340-2000	
6	NCITS 361-2002	Device Configuration Overlay (DCO)
7	INCITS 397-2005	Trzy części: Logical, Serial, Parallel ATA
8	INCITS 452-2008 INCITS 482-2012	Hybrid drive

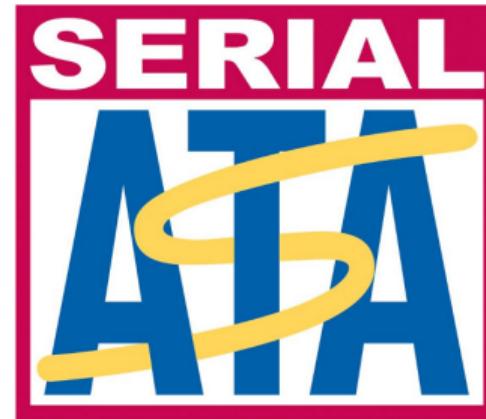
* Wycofane

- HPA i DCO — ukryte sektory, niedostępne dla użytkownika.
- Warstwa fizyczna: SATA 1.0 (1.5 GBit/s = 150 MB/s), SATA 2.0 (3 GBit/s = 300 MB/s), SATA 3.0 (6 GBit/s = 600 MB/s).

Kable ATA



Parallel ATA



Kontrolery dysków

RAM bufory we/wy

ROM zwykle EEPROM: firmware dysku; czasem tylko bootloader (HDD BIOS),
a firmware i dane są zapisane na dysku w tzw. *service area*

Sterownik sterowanie silnikami platerów (*spindle motor*)
oraz głowic (*voice coil*); zasilanie układu

Kontroler mikroporcesor, przeważnie Intel 8051, Motorola M68k, ARC 600 lub ARM 7.

Przykładowa konfiguracja HDD PCB

Pamięć RAM Hynix HY57V641620E, 8MB SDRAM 166 MHz

Pamięć ROM ST Microelectronics M29F102BB, 128kB NAND flash

Sterownik silnika ST Microelectronics L7250, 3 × 33 MHz serial

HDD Controller Infineon 0A29470



Współczesne dyski magnetyczne



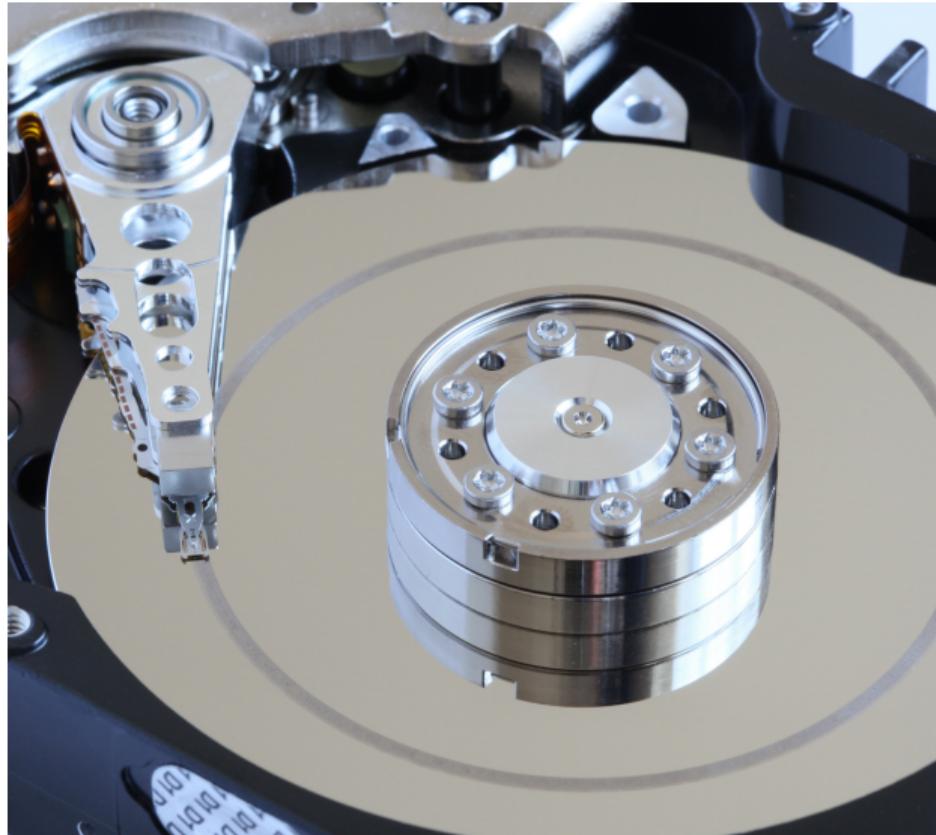
© Wikimedia Commons

Współczesne dyski magnetyczne (2)

- Wysięgniki przesuwane dawniej za pomocą silnika krokkowego, obecnie — cewki umieszczonej w polu magnesu neodymowego (*voice coil*).
- Przyspieszenie głowic dochodzi do 500 g. Czas przesuwania głowic nie przekracza 1 ms.
- Prędkość obrotowa platerów wynosi obecnie 4–15 tys. obr/min., przeważnie 5400 lub 7200 obr/min.
- Osobne głowice do zapisu i odczytu. W odróżnieniu od zapisu analogowego nie ma głowicy kasującej.
- Główice unoszą się nad powierzchnią dysku na poduszce powietrznej dzięki sile aerodynamicznej, która zależy od temperatury i gęstości powietrza — zwykle nie mogą pracować na wysokości powyżej 3000 m n.p.m.
- Dyski oddychają poprzez otwór wentylacyjny wyposażony w filtr mikrocząstek. Nowe konstrukcje mają hermetyczne nadciśnieniowe obudowy wypełnione helem.



Head crash — najczęstsza przyczyna zniszczenia dysku



Szczelina pomiędzy głowicą i powierzchnią plateru: 3 nm

Długość fali światła widzialnego: 400–700 nm

Dym tytoniowy: 10–4000 nm

Pył zawieszony: 10 nm – 10 μm

Wirusy: 5–300 nm

Ludzki włos: 17–180 μm

Otwarcie obudowy dysku poza pomieszczeniem czystym powoduje zanieczyszczenie platerów, które zwykle uniemożliwia odczyt danych.

Pomieszczenie czyste (*clean room*)



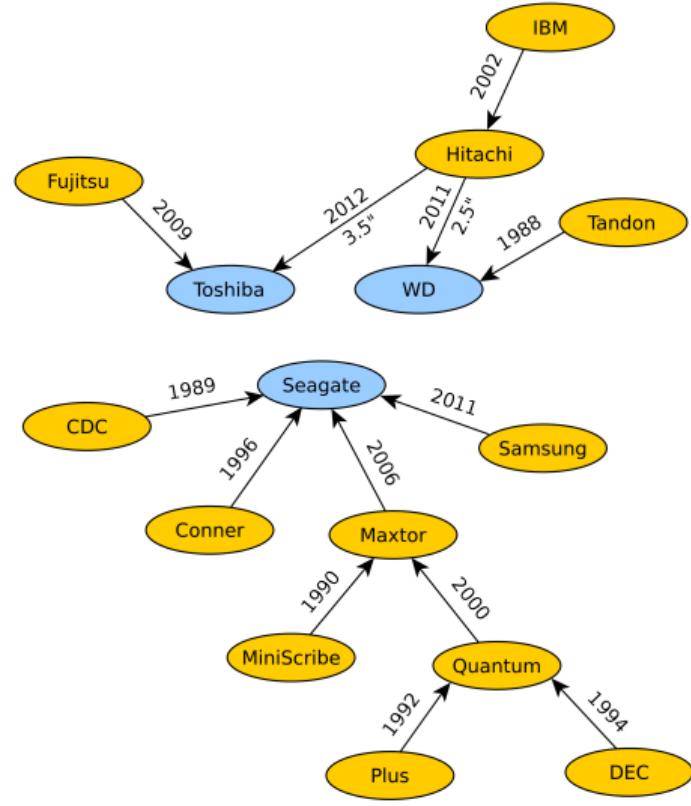
Clean room w firmie Seagate © Wikimedia Commons

Producenci dysków twardych

Obecnie trzy firmy produkują większość dysków twardych:

- Seagate
- Western Digital
- Toshiba

To jest bardzo zaawansowana technologia!



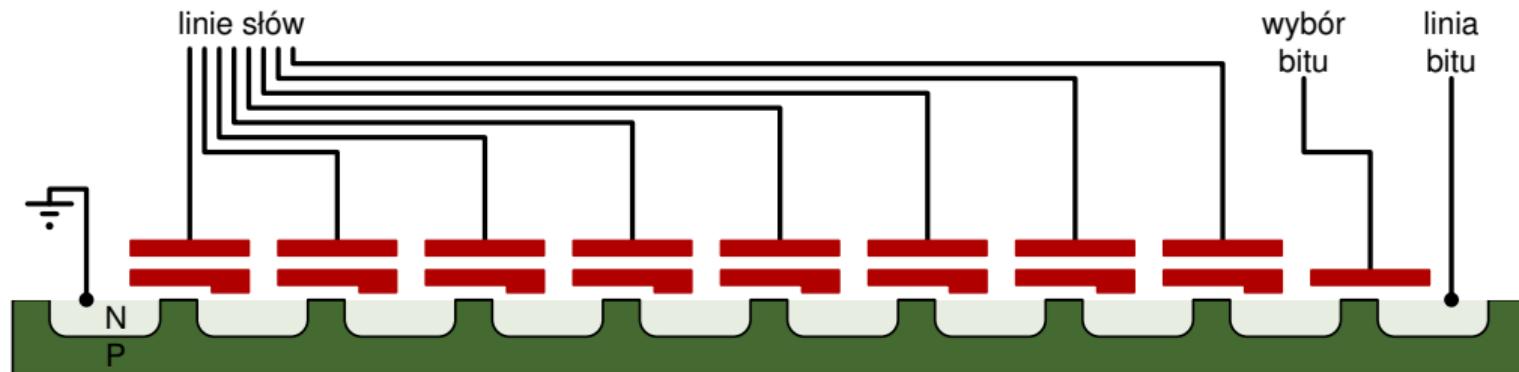
Dyski SSD



- Wykorzystują pamięci NAND flash.
- Mniejsze, lżejsze i szybsze niż dyski mechaniczne.
- Brak części ruchomych (większa niezawodność i odporność na przeciążenia).
- Typowe pojemności: 120, 240, 480, 960 GB.
- Nadal droższe niż dyski mechaniczne (w przeliczeniu na GB).

Pamięci NAND flash: zasada działania

- Tranzystory z pływającą bramką (MOSFET z dodatkową izolowaną bramką).
- Pamięć podzielona na *strony* po 512 B – 8 KiB (odpowiadające sektorom dysku).
- Strony zgrupowane w *bloki* po 16 KiB–4 MiB (32–512 stron na blok).
- Zapis (ustawienie 0): wstrzykiwanie tunelowe gorących elektronów do pływającej bramki; złapane elektrony pozostają na izolowanej bramce przez kilka lat; zapisuje się pojedyncze strony.
- Kasowanie (ustawienie 1): uwolnienie tunelowe zimnych elektronów z pływającej bramki; kasuje się całe bloki; po 100 tys. – 1 mln kasowań bramka ulega zużyciu.



Pamięci NAND flash

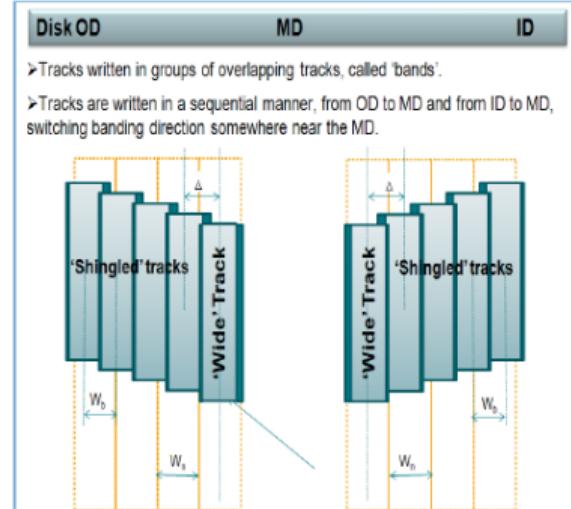
- *Mapa stron*: modyfikacja strony polega na zapisaniu jej nowej wartości w inne miejsce (bo kasowanie odbywa się blokami po 32–512 stron).
- Inaczej niż w dyskach twardych strony o kolejnych numerach nie muszą sąsiadować.
- Bloki z których wszystkie strony zostały przeniesione zeruje się i zwraca do puli wolnych bloków.
- *Wear levelling*.
- Jeśli mało wolnych bloków, to następuje *write amplification*.
- Dobra jeśli dysk wie, które strony nie są używane: *trimming*.

Shingled Magnetic Recording



Takeaways

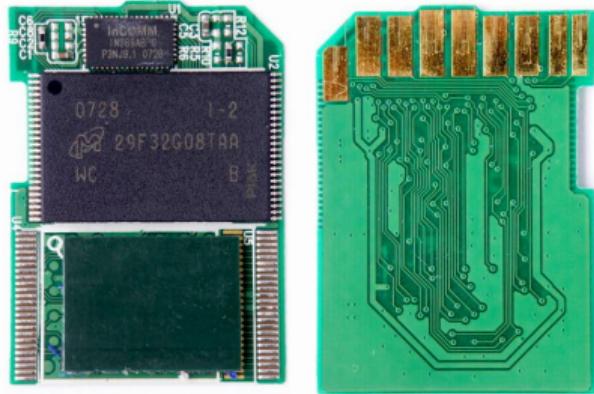
- SMR- increase capacity over conventional recording
- Allows more narrow tracks to be written
- Performance transparent in the vast majority of use cases



Dyski SMR przypominają dyski SSD

- Sektor fizyczny dysku magnetycznego: zwykle 4 kiB (Advanced Format).
- Strona zapisu pamięci flash: 4–8 kiB.
- Strefa zapisu dysków SMR: 20–40 MB.
- SSD erase blocks: zwykle 128–512 stron, tj. 512 kiB – 4 MiB.
- Cache CMR dysku SMR: rzędu 25 GB.
- Cel dysków SMR: osiągnąć gęstość zapisu ponad 1 Tb/in².
- SMR często jako dyski hybrydowe (duża ilość DRAM i NAND flash).

Bunnie & Xobs: *On Hacking MicroSD Cards*, 30C3 (2013)



- Procesor: przeważnie 8051 lub ARM, zegar rzędu 100 MHz.
- Można obniżyć koszty i produkować pamięci z defektami — mapa stron będzie pomijać uszkodzenia. Nawet 80% stron może być uszkodzonych.
- Firmware daje się aktualizować!
- Zastosowanie: Arduino (8-bit, 16 MHz) kosztuje \$20, karta SD mniej niż \$1, a dodatkowo ma pamięć masową rzędu kilku gigabajtów!
- Pendrive działają podobnie.
- Uwaga na wirusy w firmware!

Archeologia cyfrowa

Christopher Fenton, *Digital Archeology with Drive-Independent Data Recovery: Now, With More Drive Dependence!* ELEN E9002 Research Project Final Report, Summer 2011.

