



Ukladanie údajov

HDD · SSD · Rozhrania · Spôsoby záznamu

Predmet: Počítačové Architektúry

Autor: Tobiáš Lehen

Univerzita Cyrila a Metoda | Akademický rok 2025/2026

Obsah prezentácie

01

Úvod – pamäťová hierarchia

Von Neumann, princípy lokality, IOPS vs. latencia

02

HDD – mechanická architektúra

Platne, hlavy, aktuátor, spindle, geometria CHS/LBA

03

HDD – technológie záznamu

PMR, SMR, HAMR (Mozaic 3+), MAMR, areal density

04

SSD – NAND Flash architektúra

SLC/MLC/TLC/QLC/PLC, 3D NAND, 321–400+ vrstiev

05

SSD – Kontrolér, FTL, spoľahlivosť

FTL, Wear Leveling, Garbage Collection, ECC/LDPC

06

Rozhrania – SATA, SAS, NVMe, PCIe

Protokoly, šírka pásma, PCIe Gen5/6, CXL 3.0

07

Spôsoby záznamu a kódovania

RLL, PRML, WAF, ZNS, FDP, ECC

08

Výkonnostné porovnanie & Trh

Benchmarky, IOPS, trhové štatistiky 2024–2025

Úvod – paměťová hierarchia

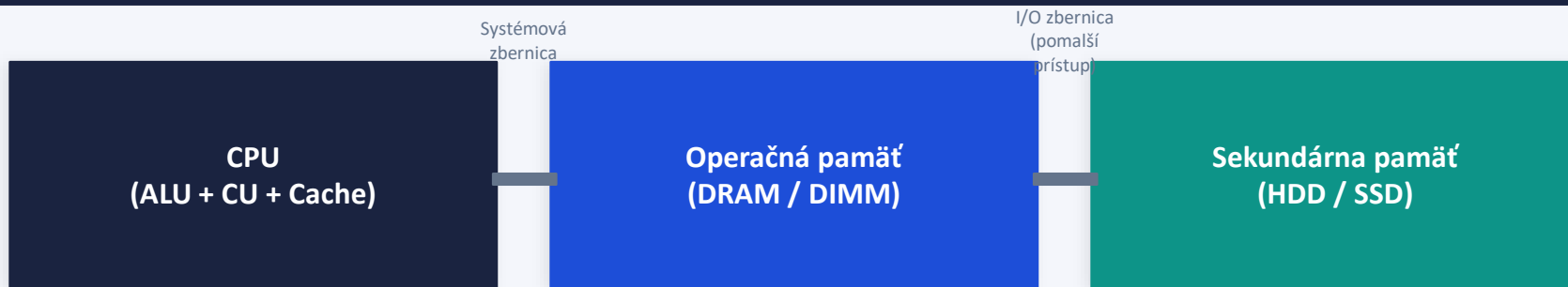
Von Neumann · principy lokality · kompromizy rychlost/kapacita/cena

01

Paměťová hierarchie – Memory Hierarchy



Architektúra a rola sekundárnej pamäte



Volatilita RAM

DRAM je volatilná – po vypnutí napájania stráca všetky dáta. Sekundárna pamäť (HDD/SSD) uchováva dáta perzistentne. Toto je fundamentálny dôvod existencie sekundárnej pamäte v Von Neumannovej architektúre.

Princíp lokality

Temporálna lokalita: nedávno použité dáta sa pravdepodobne opäť použijú. Priestorová lokalita: dáta blízko aktuálne použitých dát sa tiež pravdepodobne použijú. OS využíva tieto princípy na prefetching a page cache.

Hierarchia & bottleneck

CPU nie je schopné pristupovať priamo k disku. Dáta putujú: HDD/SSD → RAM → L3 Cache → L2 → L1 → Register. Storage latencia (ms) vs. CPU cyklus (ns) = 1 000 000x rozdiel – storage je najväčší bottleneck systémov.

HDD – Mechanická architektúra

Platne · čítacie hlavy · aktuátor · geometria CHS/LBA · výkon

02

HDD – Fyzická architektúra a kľúčové komponenty

Platne (Platters)

Rotujúce disky z hliníka alebo skla potiahnuté feromagnetickým materiálom (CoFeB/FePt). Moderné high-cap HDD: 7–10 platní v heli-om-naplnenej komore (Seagate Exos, WD Gold). Helium znižuje trenie a turbulencie.

Čítacia/Zapisovacia hlava

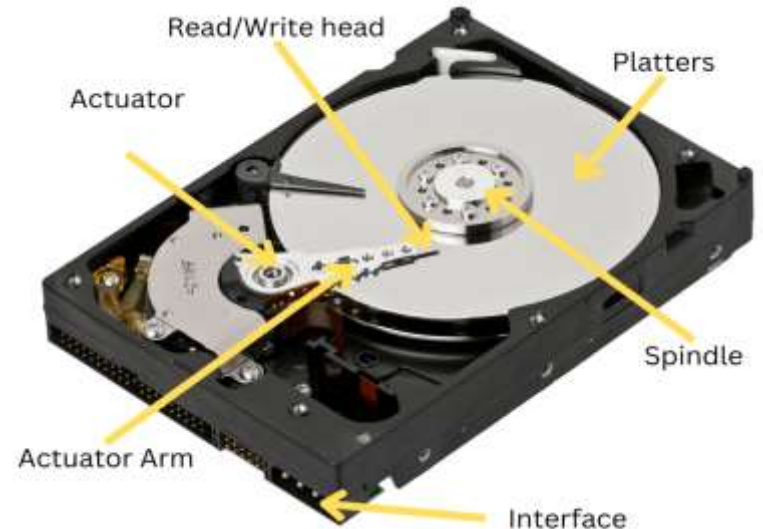
Indukčná zapisovacia cievka + magnetorezistívny (MR) senzor na čítanie. Vzduchová medzera $\sim 1\text{--}3\text{ nm}$ (tzv. head fly height). Gen 7 Spintronic čítacie hlavy (Seagate Mozaic 3+) s ultra-nízkym šumom.

Aktuátor – VCM

Voice Coil Motor (VCM) presúva rameno s hlavami pomocou elektromagnetickej sily. Rotačný pohyb cez kalibrovaný pivoť. Dual-actuator HDDs (Western Digital UltraSMR): dva nezávislé VCM na jednom disku $\rightarrow$ dvojnásobný IOPS.

Spindle Motor (BLDC)

Bezkomutátorový elektromotor otáča platňami: 5 400 RPM (NAS, archív), 7 200 RPM (desktop/enterprise), 10 000/15 000 RPM (legacy SCSI/SAS). Rýchlosť platní priamo určuje rotačné oneskorenie (avg. 4 ms pri 7200 RPM).



HDD – Adresovanie: Stopy, Sektory, Cylindre, LBA



Track (Stopa)

Sústredná kružnica na povrchu platne. Hustota: 600 000+ stôp/palec (TPI). Vonkajšie stopy sú dlhšie → ZBR (Zone Bit Recording) ukladá viac sektorov/stopu.



Sektor

Základná adresovateľná jednotka. Legacy: 512 B (512n), moderné: 4 096 B (Advanced Format / 4K). Fyzický sektor 4K + emulovaný 512B = typ 512e (dnes bežný). Štruktúra: header | dáta | ECC.



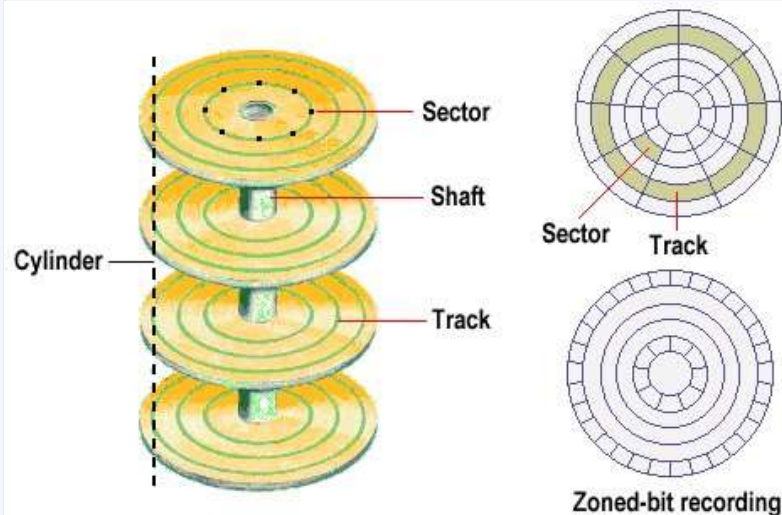
Cylinder

Všetky stopy na rovnakom polomere naprieč všetkými platňami. Čítanie z celého valca bez pohybu ramena (seek = 0) → optimalizácia výkonu pri sekvenčnom prístupe.



CHS → LBA48

Cylinder-Head-Sector adresovanie (max 8 GB) nahradené LBA. LBA48 (48-bitové LBA) adresuje $2^{48} \times 512 \text{ B} = 128 \text{ PiB}$. LBA64 pre budúce 4K-native zariadenia nad 128 PiB.



Výkonnostné parametre HDD

Seek time (avg.)

3–8 ms (7 200 RPM)

Rotačné oneskorenie

4.2 ms (7 200 RPM)

Max. transfer

~300 MB/s (outer track)

HDD – Technológie záznamu

PMR · SMR · HAMR (Mozaic 3+) · MAMR · areal density roadmap

03

Technológie magnetického záznamu – PMR, SMR, HAMR, MAMR

PMR – Perpendicular MR

Bity orientované kolmo na platňu (od ~2005). Štandard pre ~1–3 TB/plaťňu. Koercivita 10–20 kOe. Stále používané v mid-range HDD. Limit hustoty: ~1.5 Tbit/in².

SMR – Shingled MR

Stopy sa prekrývajú ako škridle (20–25 % väčšia hustota). Výhoda: kapacita za nízku cenu. Nevýhoda: pomalý náhodný zápis (write-back penalty). WD Red (archív), Seagate IronWolf.

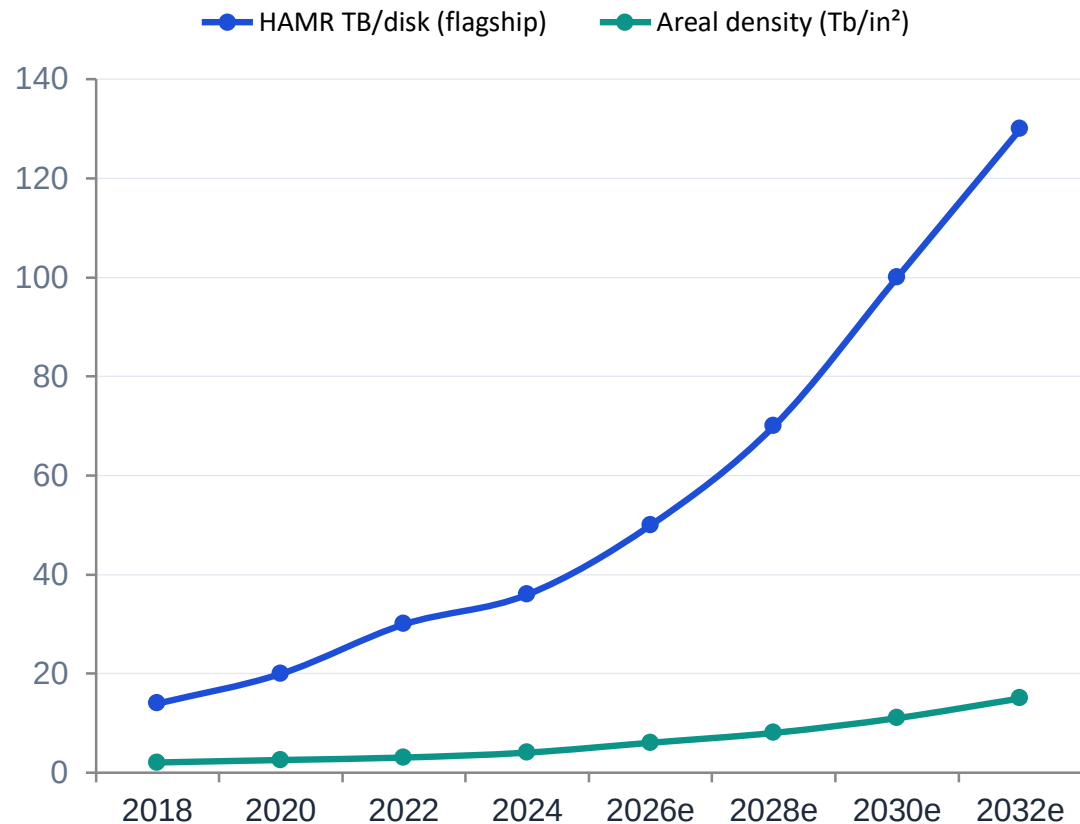
HAMR – Heat-Assisted MR

Laser (790 nm, <1 mW) zahreje bod na platni na ~450 °C → znižuje koercivitu FePt media → umožní zápis hustejších bitov. Seagate Mozaic 3+: 3 TB/plaťňu, prod. 2024. Demo: 6.5 TB/plaťňu v 2024. Roadmapa: 100 TB/disk do 2032.

MAMR – Microwave-Assisted MR

Spin-Torque Oscillator (STO) generuje mikrovlnné pole, ktoré znižuje koercivitu bez zahrievania. WD MAMR: 26–32 TB HDDs (2024). Energeticky efektívnejší ako HAMR. Areal density ~1.5–2 TB/plaťňu.

HDD Kapacitná roadmapa – Seagate Mozaic 3+ / 4+ / 5+



Zdroj: Seagate – The nanomaterials enabling the AI era (2025) | TechRadar – HAMR and SMR: 6TB/platter demo H1 2025 | Tom's Hardware – Seagate Mozaic roadmap

2024

Mozaic 3+: 30 TB HDD (HAMR, 3 TB/platňu, 10 platní). Demo: 6.5 TB/platňu v laboratóriu (H1 2025).

2025

Seagate dodal >550 EB kapacity HAMR diskov. >1 mil. HAMR kusov v prevádzke u cloudových zákazníkov.

2026

Mozaic 4+ (kvalifikácia): 4 TB+/platňu, cieľ 40 TB+ na disk. 2× energetická efektívnosť oproti PMR.

2026e

WD HAMR: 40–50 TB (14-platter design). Seagate Mozaic 4+ v plnej sériovej výrobe.

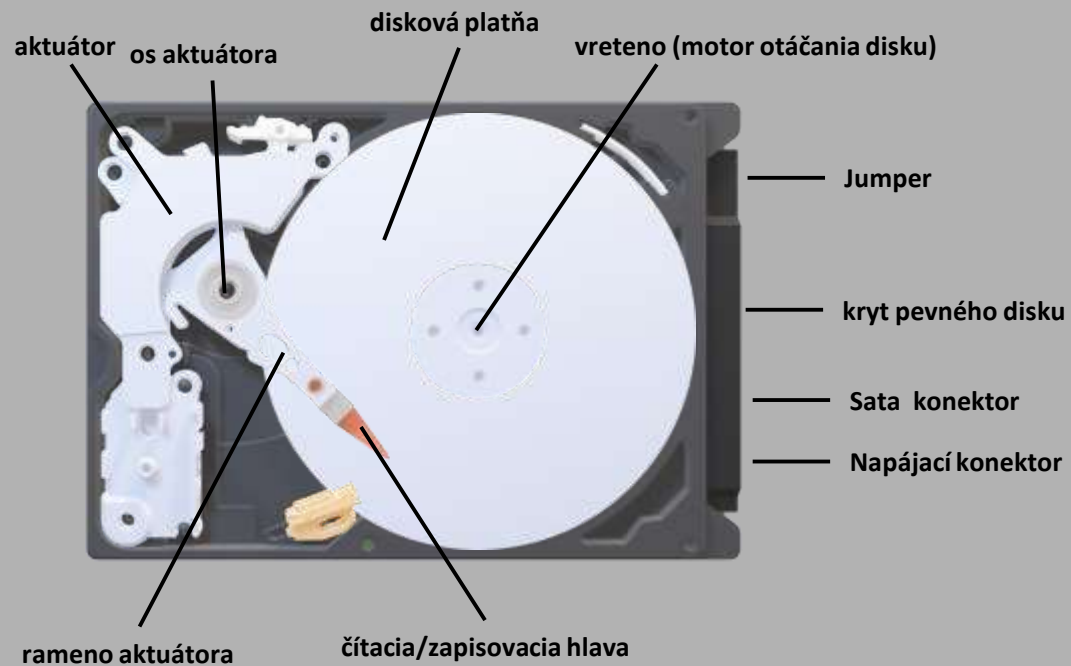
2028e

Seagate demonštruje 10 TB/platňu (Mozaic 5+ vývoj). Cieľ: 60–80 TB/disk.

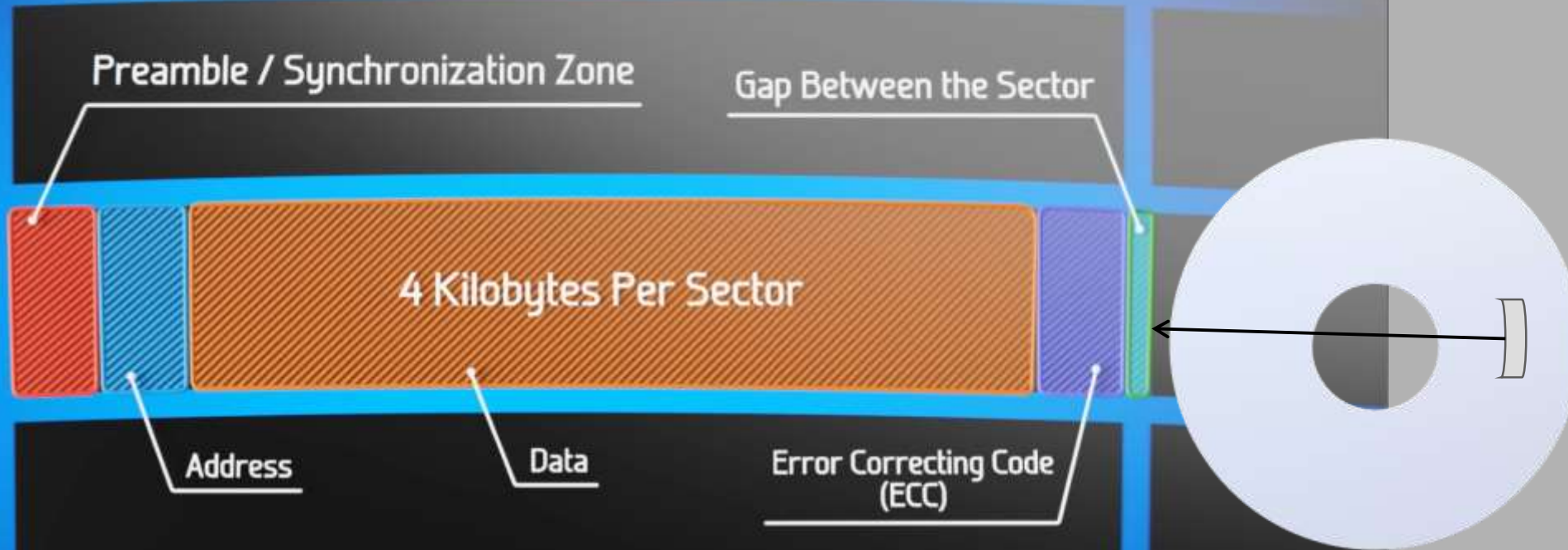
2032e

Roadmapa k 100 TB+/disk. FePt media + bit-patterned media ako dlhodobá perspektíva.

HDD komponenty

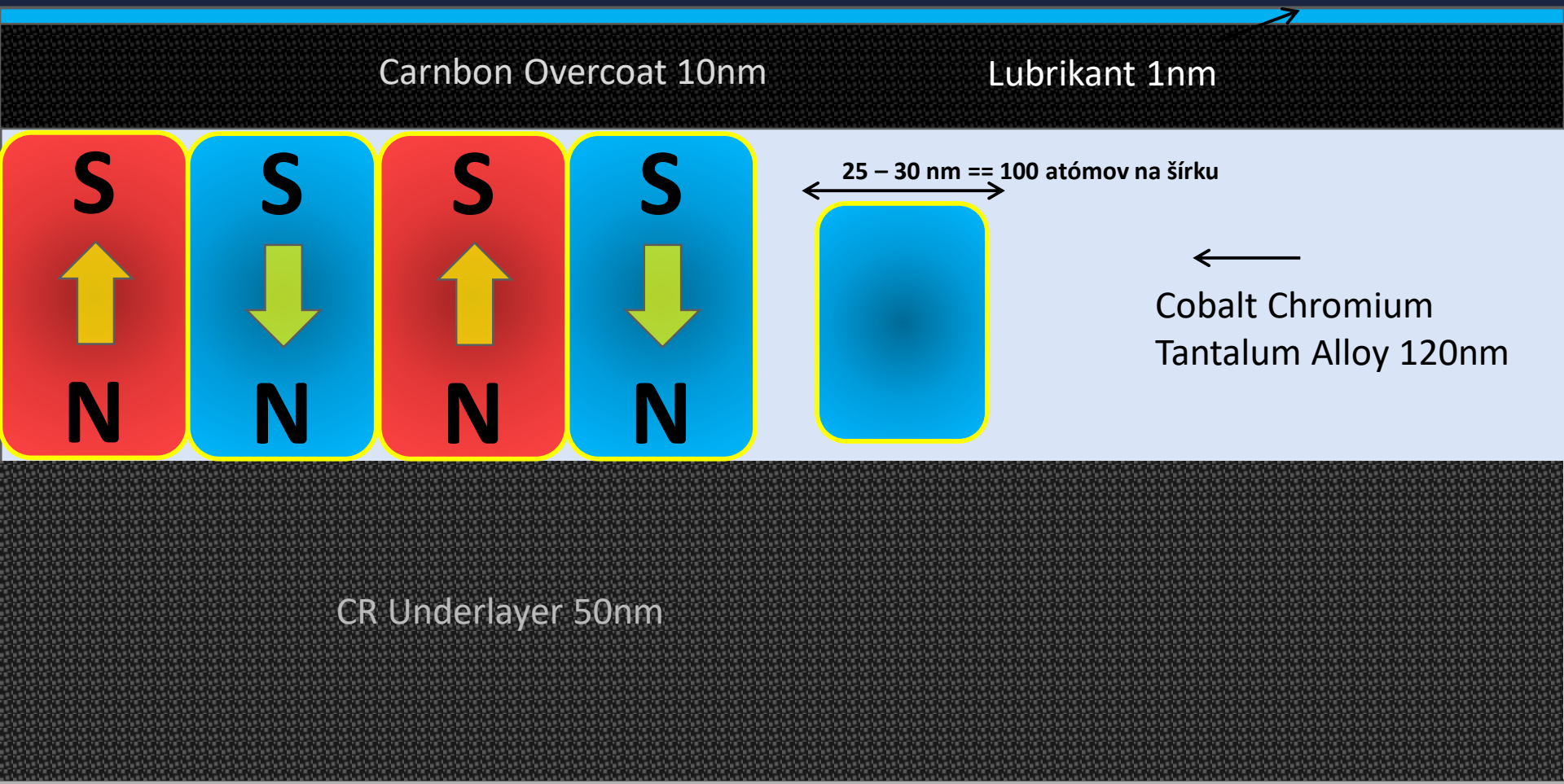


HDD komponenty

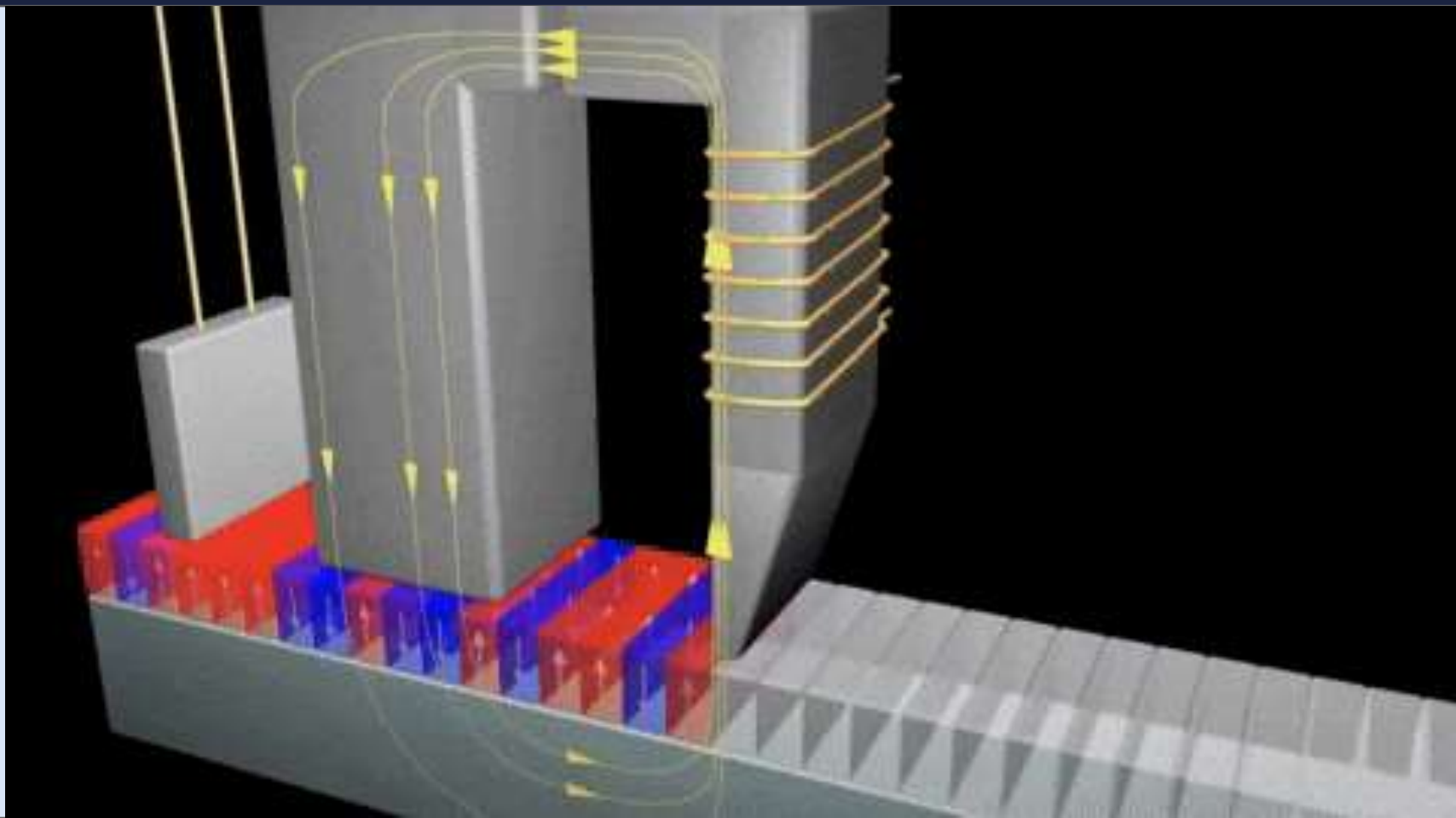


HDD komponenty

1 000 000 + zoom



HDD komponenty



SSD – NAND Flash architektúra

Bunky · stránky · bloky · 3D NAND

04

NAND Flash – Typy buniek: SLC / MLC / TLC / QLC / PLC

SLC – 1 bit/cell

~100 000 P/E cyklov

Najrýchlejší a najtrvanlivejší typ. Cena: ~5× TLC. Použitie: enterprise SSD cache, NVDIMM, AI SSD (SK Hynix + NVIDIA: cieľ 100M IOPS/SSD do 2027). Napätové hladiny: 2.

MLC – 2 bity/cell

~10 000 P/E cyklov

Vyvážený výkon a trvanlivosť. Cena: ~3× TLC. Dnes prevažne v enterprise NVMe SSD (Intel/Solidigm P5510, Samsung PM9A3). Napätové hladiny: 4.

TLC – 3 bity/cell

~1 000–3 000 P/E cyklov

Dominantný typ od ~2018. Samsung 990 PRO, WD Black SN850X, SK Hynix Platinum P41. Napätové hladiny: 8. SLC cache kompenzuje nižší zápis na prázdnych blokoch.

QLC – 4 bity/cell

~100–1 000 P/E cyklov

Nízka cena, vysoká kapacita. Samsung 870 QVO, Micron 6550 ION (120 TB, 2025), Solidigm P41. Nástup v dátových centrách pre cold/warm storage tier. Napätové hladiny: 16.

PLC – 5 bitov/cell

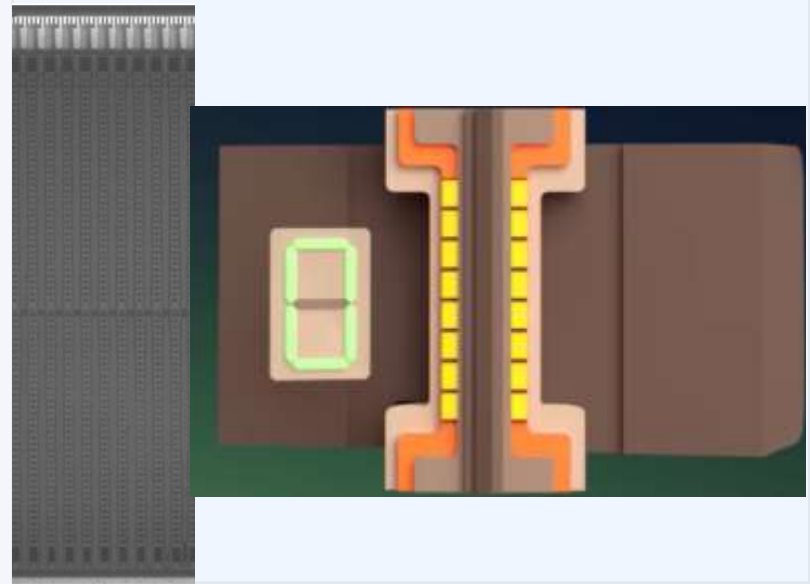
<100 P/E cyklov (exp.)

Experimentálna fáza. Intel/Solidigm výskum. Hustota +20% vs QLC. Limit: obrovský počet napätových hladín (32) → extrémna citlivosť na šum. Výhradne pre studený archív.

3D NAND – Vertikálna architektúra a závodenie vo vrstvách (2024–2025)

Kľúčové fakty 3D NAND

- 2D planar NAND dosiahla fyzikálny limit pri ~15–16 nm procesu (~2014).
- 3D NAND (Samsung V-NAND od 2013) skladá bunky vertikálne → menej závislé na litografii.
- Charge Trap Flash (CTF) dominuje – elektrony zachytené v nitridovej vrstve, nie v FG.
- Hybrid Bonding: periféria a bunky na separátnych wafre, spojených Cu-Cu bondom → SK Hynix 321-layer QLC (mass prod. 2025).
- Samsung V-NAND 9. gen: 290 vrstiev TLC, masová výroba 2024. Cieľ: 1 000 vrstiev do 2030.
- Micron 2650 SSD: 276-layer 3D NAND, júl 2024. SK Hynix vyvíja 400+ vrstiev (plán: 2026).
- Kapacita Die: ~1–2 Tb (TLC), až ~4 Tb (QLC) pri najnovších generáciách.



Generácie 2024–2025

Samsung (2024)	290 vrstiev TLC
Micron 2650 (2024)	276 vrstiev
SK Hynix 321L (2025)	321 vrstiev QLC
SK Hynix 400L (2026e)	~400 vrstiev, hybrid bond

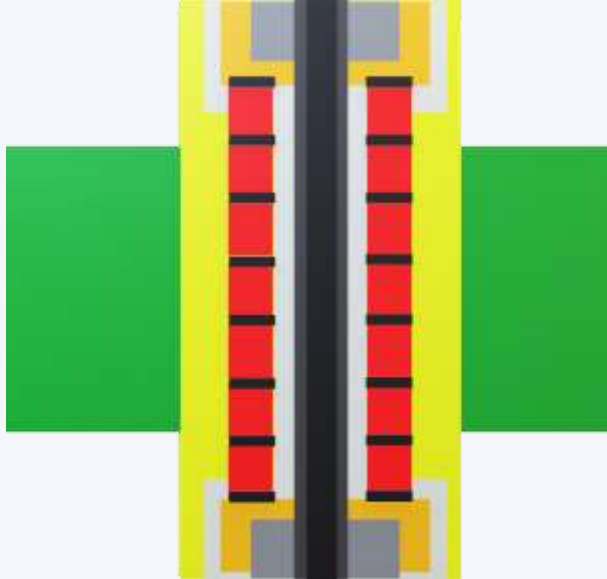
Charge Trap Flash (CTF)

Charge Trap Flash (CTF)

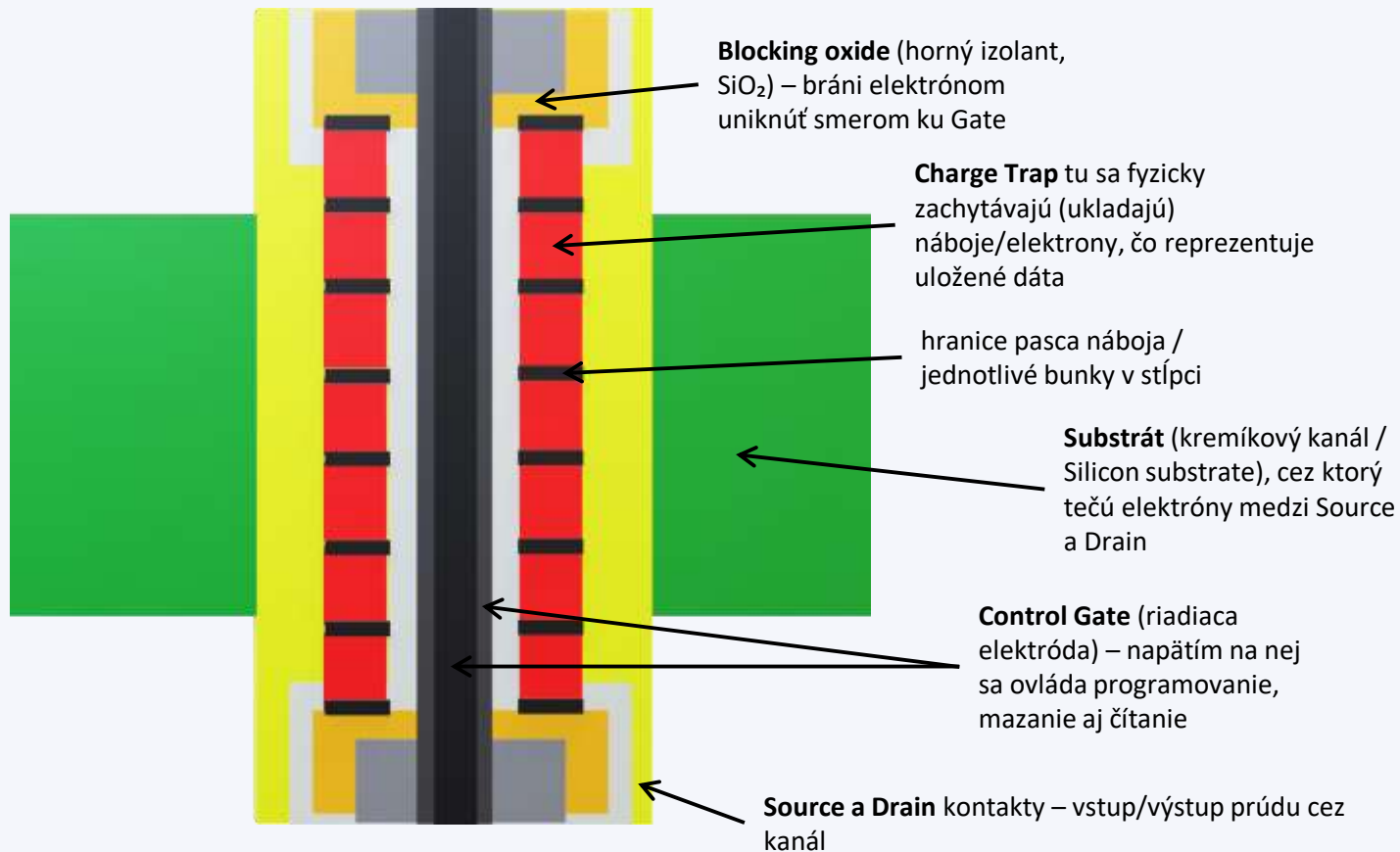
Charge Trap Flash je typ pamäťovej bunky použíwanej v moderných SSD diskoch. Každá bunka obsahuje **izolačnú vrstvu z nitridu kremíka (Si_3N_4)**, ktorá slúži ako pasca pre elektrické náboje.

Pri **zápise** dat sa priloží vysoké napätie, ktoré vytlačí elektróny cez tenkú oxidovú vrstvu do nitridu, kde sa zachytia. Pri **čítaní** sa zmeria elektrické napätie bunky – prítomnosť alebo neprítomnosť zachytených elektrónov reprezentuje hodnotu **0 alebo 1**. Pri **výmaze** sa priloží obrátené napätie, ktoré elektróny uvoľní a bunku resetuje do pôvodného stavu.

Vďaka tejto jednoduchšej štruktúre je možné vrstvy skladať na seba, čo je základ moderných **3D NAND** čipov s vysokou kapacitou.

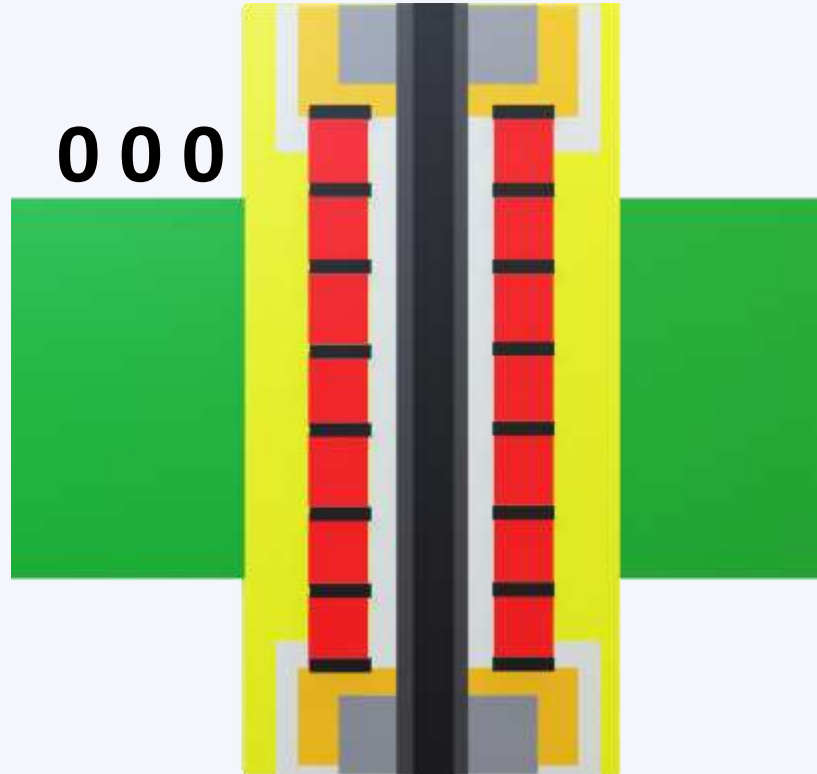


Charge Trap Flash (CTF)

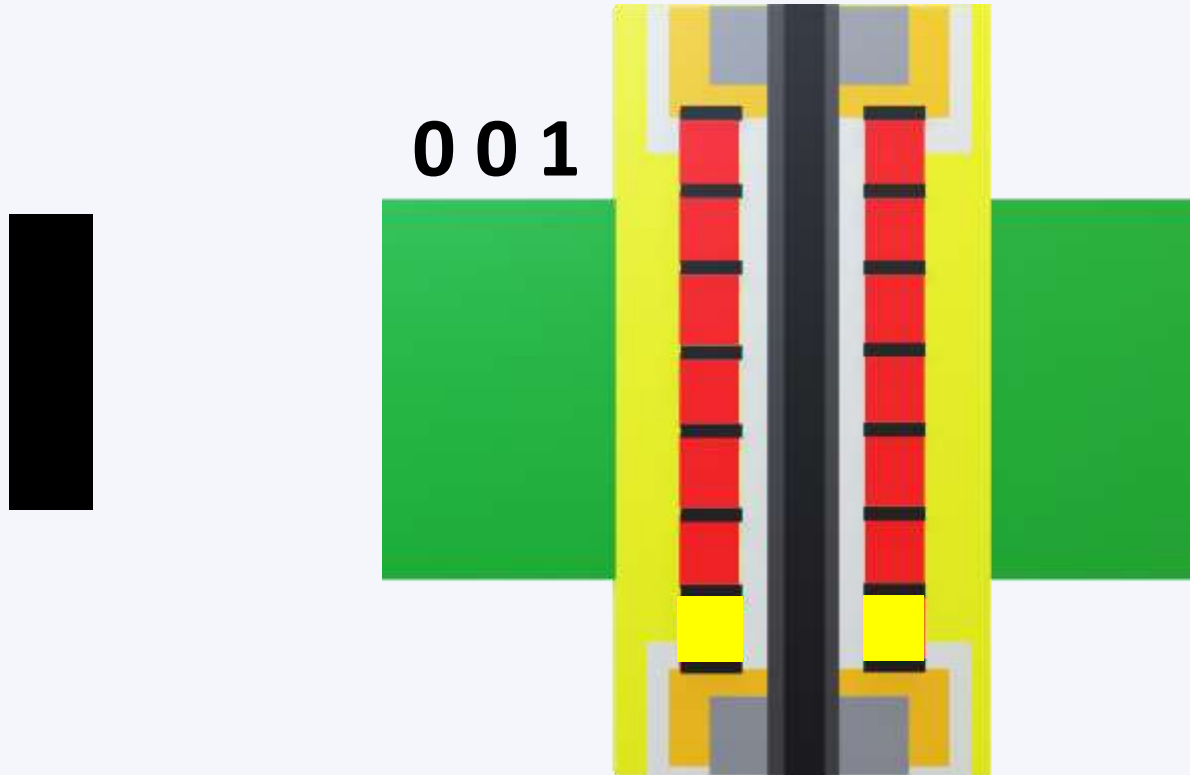


Charge Trap Flash (CTF)

0

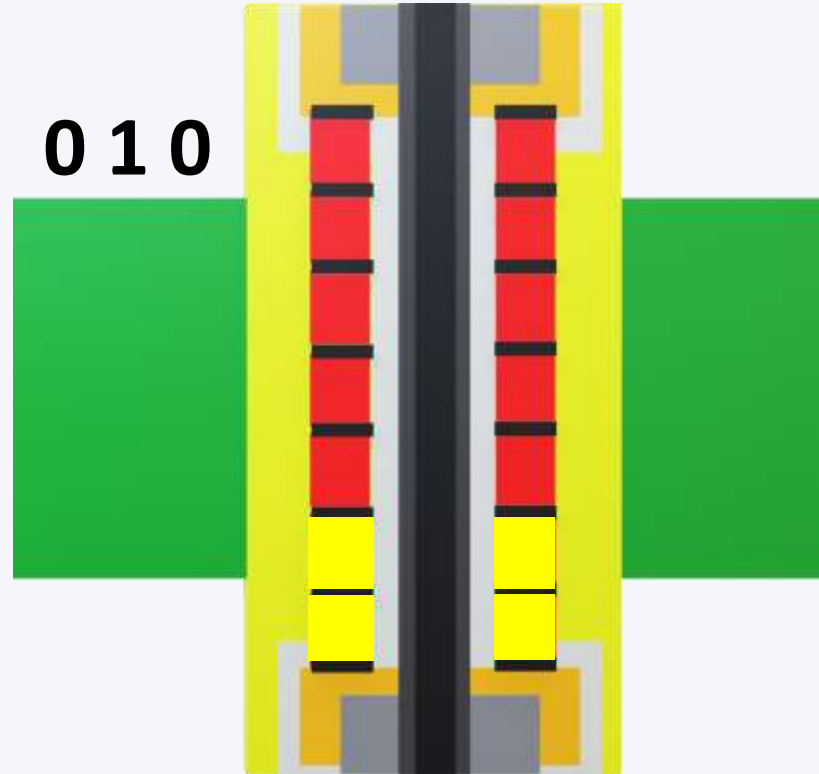


Charge Trap Flash (CTF)

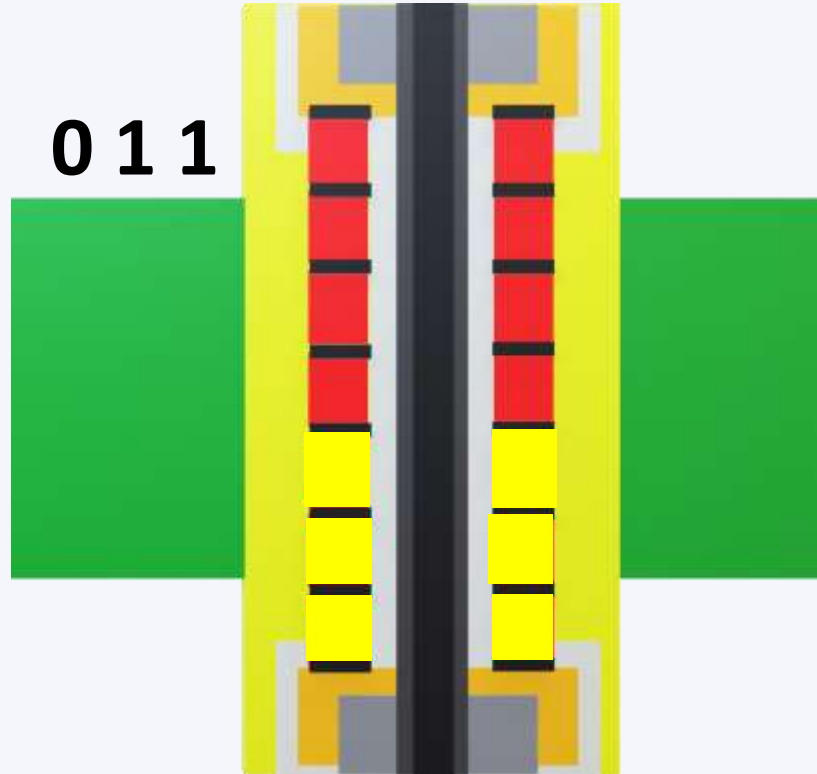


Charge Trap Flash (CTF)

2

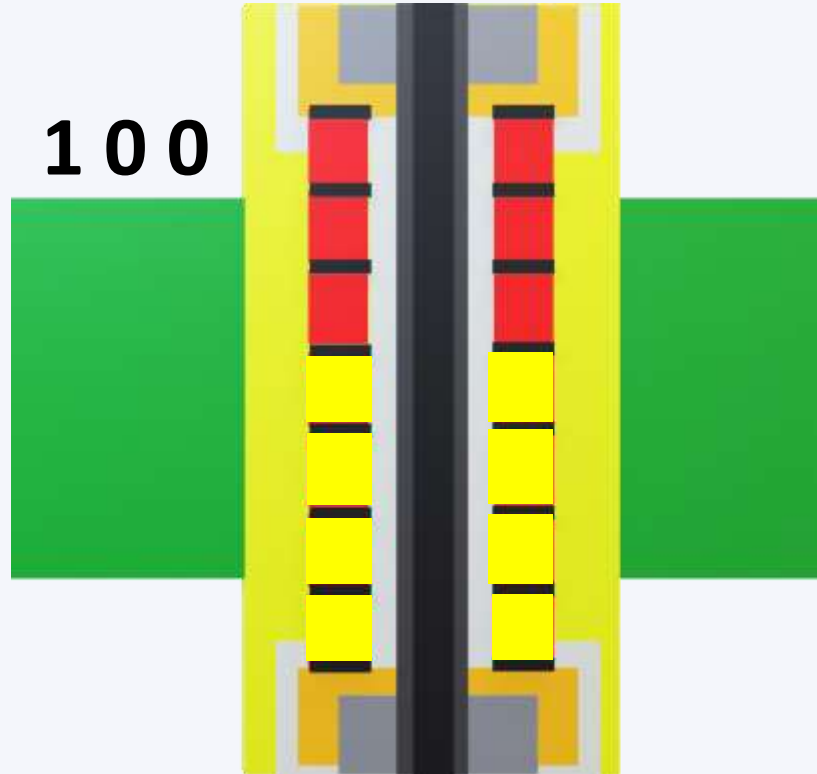


3

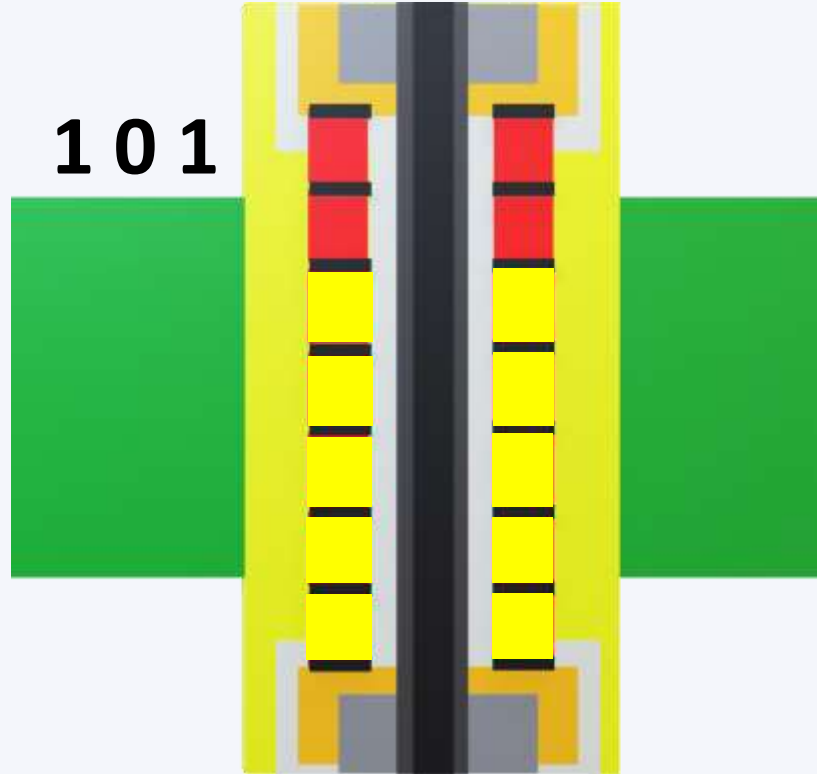


Charge Trap Flash (CTF)

4

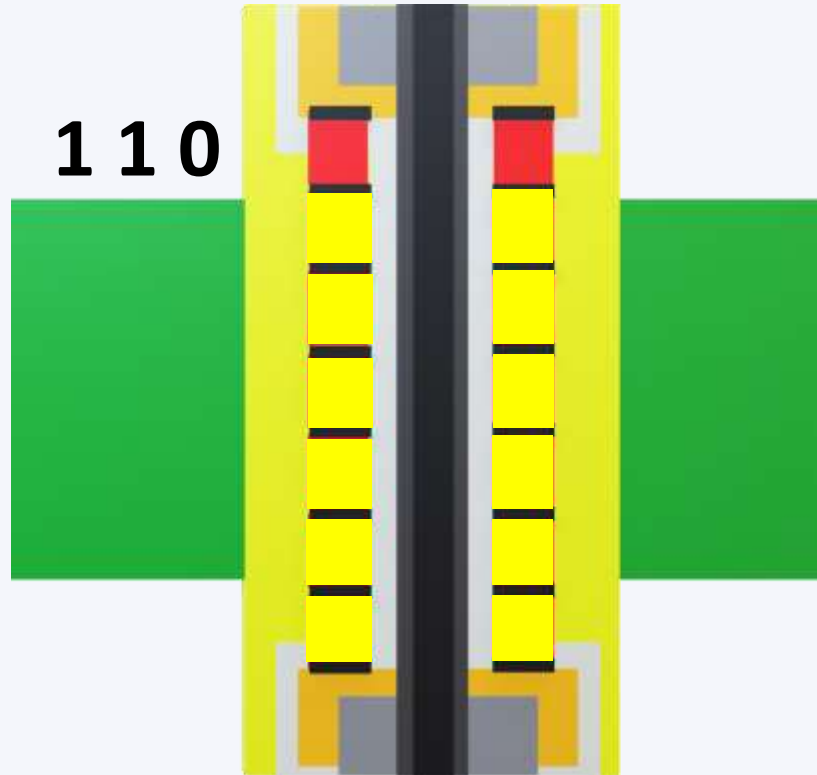


5



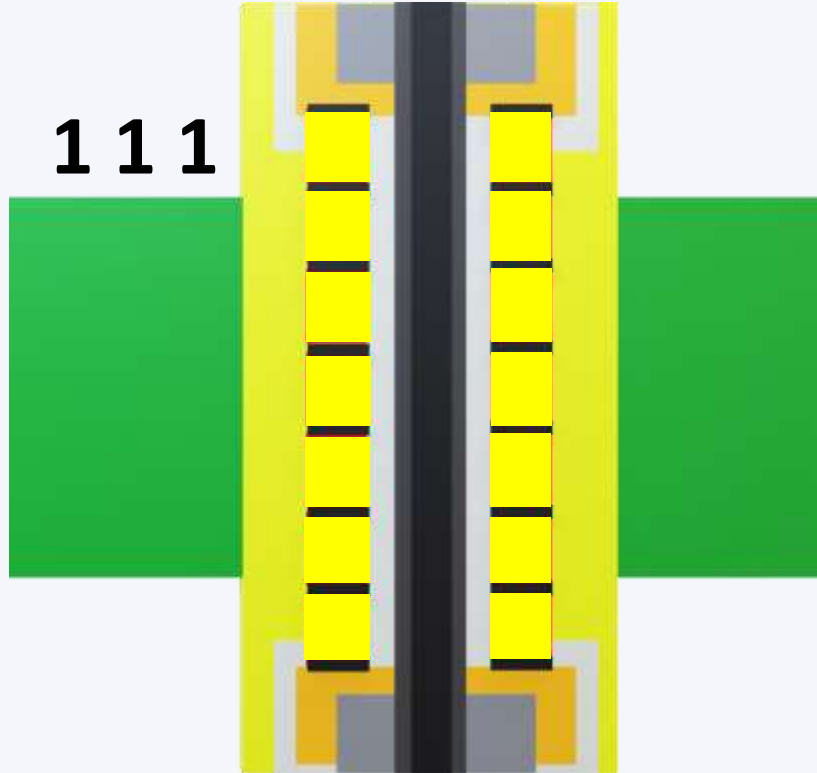
Charge Trap Flash (CTF)

6

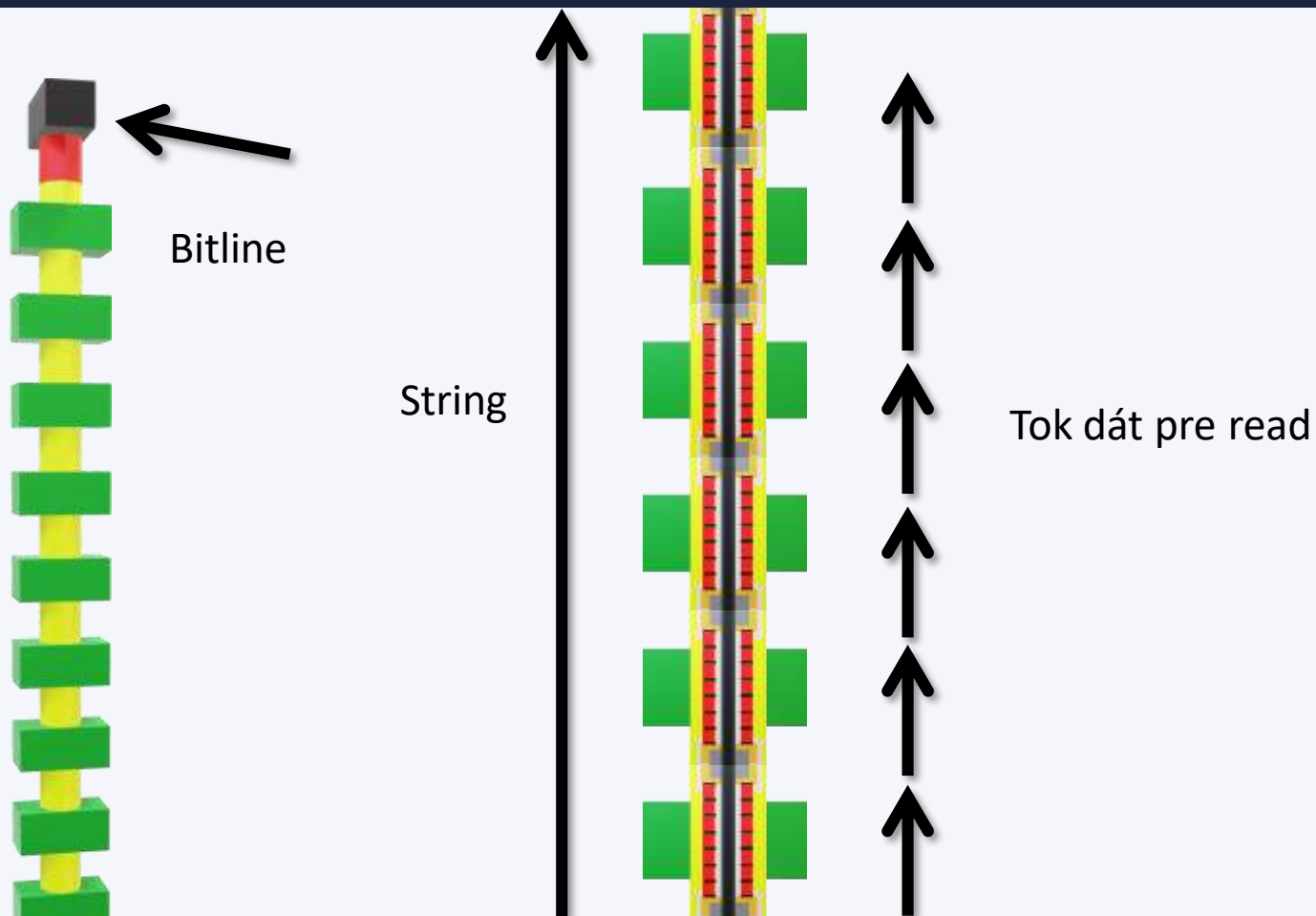


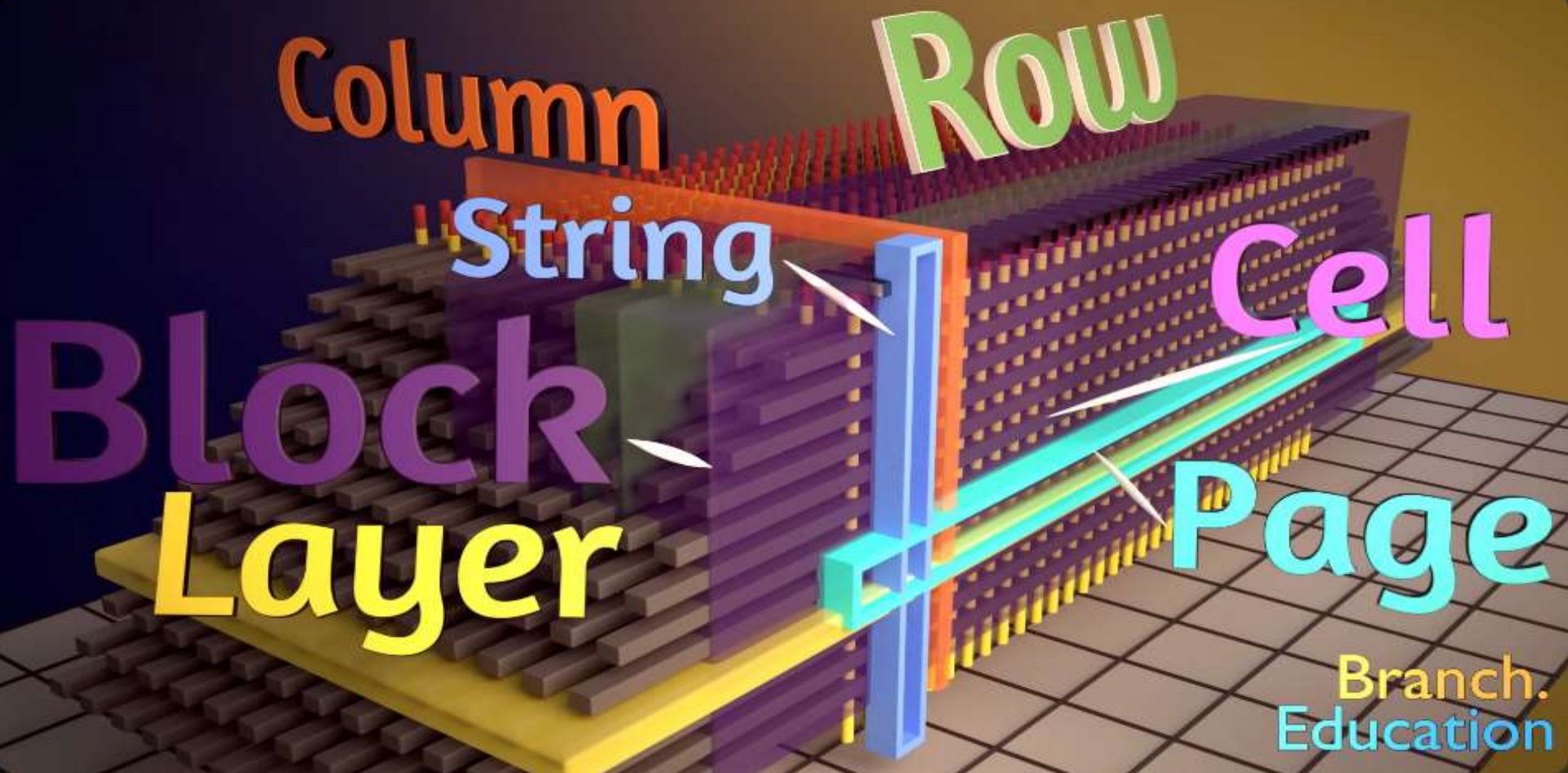
Charge Trap Flash (CTF)

7

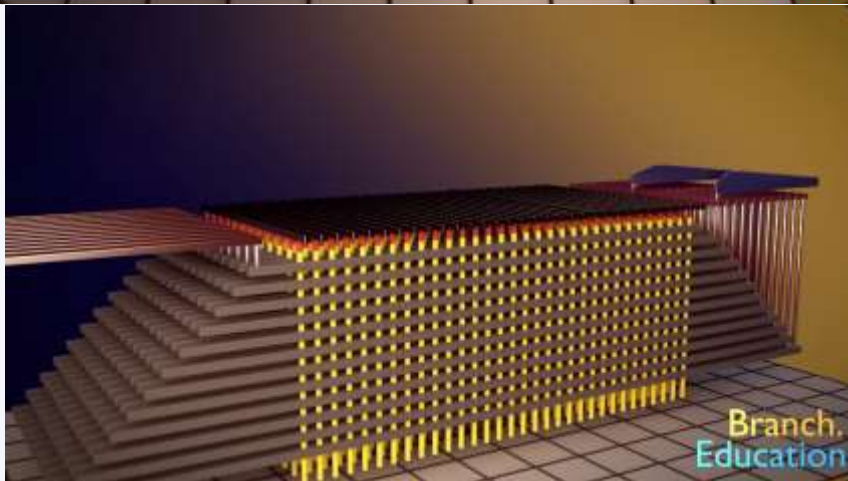
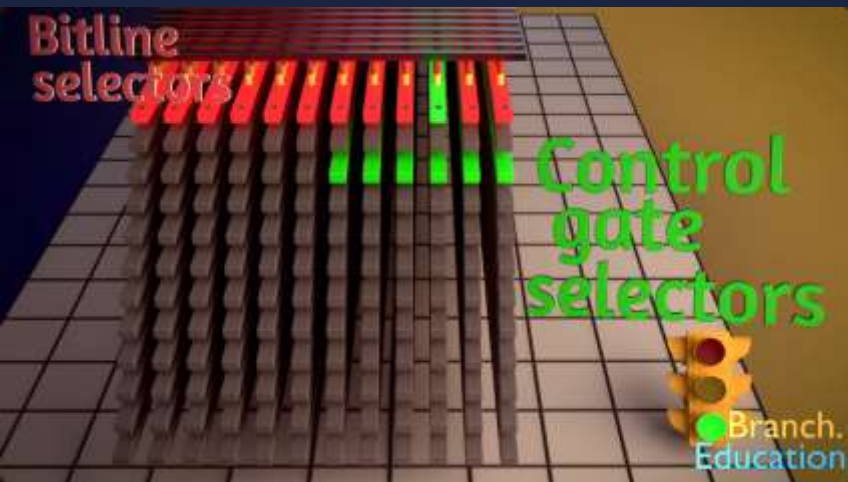


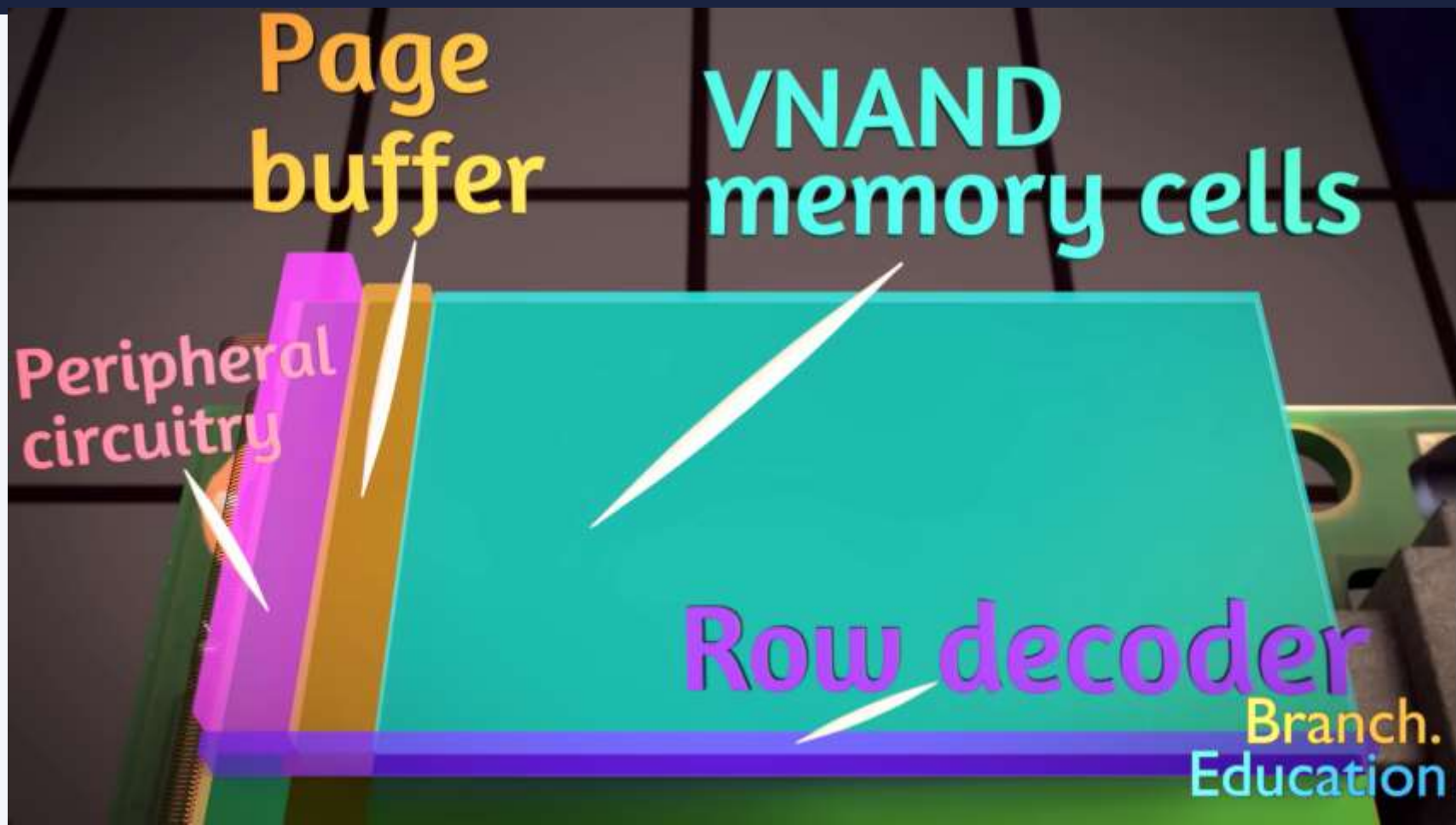
Charge Trap Flash (CTF)





ssd





SSD – Kontrolér, FTL a spoľahlivosť

FTL · Wear Leveling · Garbage Collection · ECC/LDPC · WAF

05

SSD – Kontrolér a Flash Translation Layer (FTL)

FTL – Flash Translation Layer

Mapuje LBA (logické bloky) na fyzické stránky NAND. Skrýva pred OS asymetriu: čítanie (mikros.) $\neq$ zápis stránky $\neq$ mazanie bloku (blok = 256–512 stránok). Kľúčová vrstva pre výkon a životnosť SSD.

Wear Leveling

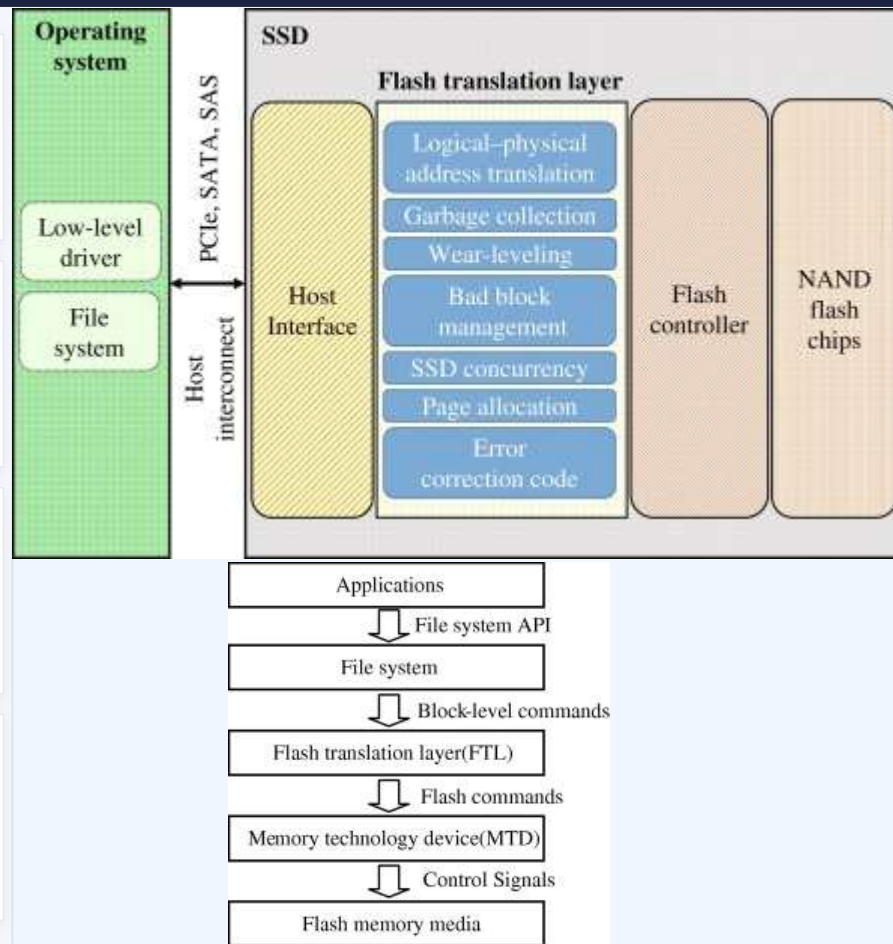
Distribuuje zápisy rovnomerne cez všetky bloky. Dynamický WL: rotuje aktívne bloky. Statický WL: presúva aj studené (málo-písané) dáta. Cieľ: vyrovnať počet P/E cyklov naprieč celou NAND.

Garbage Collection (GC)

Zbiera platné stránky z čiastočne-platných blokov do nových blokov $\rightarrow$ maže prázdne. Spúšťa sa na pozadí (idle) alebo pri nedostatku voľného priestoru (forced GC). Hlavný dôvod Write Amplification (WAF > 1).

ECC – Error Correction Code

LDPC (Low-Density Parity-Check): súčasný štandard. BCH kódy v starších generáciách. LDPC opravuje viacbitové chyby/stránku – kritické pri QLC ($>10^4$ -2 raw BER). Phison E26 LDPC engine pre PCIe 5.0 NVMe.



Write Amplification Factor, ZNS a Flexible Data Placement (NVMe 2.0)

WAF = Dáta zapísané na NAND ÷ Dáta zapísané hostiteľom (OS)

Ideálny WAF = 1.0. Reálne pri TLC consumer SSD: 1.05–3.0. Garbage Collection a Wear Leveling sú hlavné príčiny WAF > 1. Vysoký WAF skraca životnosť (TBW).

Max. P/E cykly podľa typu bunky



ZNS + FDP – Architektúra záznamu

- ZNS (Zoned Namespaces) – NVMe 2.0 funkcia: rozdeľuje SSD na zóny (~96–512 MB)
- Každá zóna sekvenčne zapisovateľná → eliminuje väčšinu GC → WAF ≈ 1.0–1.1
- FDP (Flexible Data Placement) – NVMe 2.0 doplnok ZNS: host riadi umiestnenie
- Samsung 990 Pro / WD SN850X podporujú FDP (firmvér update 2024)
- Meta a Google nasadili ZNS v RocksDB → -40% write amplification
- Linux kernel 5.9+: native ZNS podpora (f2fs, ZenFS, btrfs)
- Výsledok: predvídateľná latencia, nižšie tepelné zaťaženie, dlhšia životnosť

Rozhrania – SATA, SAS, NVMe, PCIe

Protokoly · šířka pásma · PCIe Gen5 v produkci · CXL 3.0 · M.2/U.2/EDSFF

06

Rozhrania úložísk – prehľad, šírka pásma a použitie (2025)

SATA III

Šírka pásma:	600 MB/s
Latencia:	~100 µs
Použitie:	Spotrebiteľ, legacy NAS

SAS-3 / SAS-4

Šírka pásma:	1 200/2 400 MB/s
Latencia:	~50 µs
Použitie:	Enterprise server (HDDs+SSDs)

PCIe 3.0 ×4 / NVMe

Šírka pásma:	3 940 MB/s
Latencia:	~20 µs
Použitie:	NVMe SSD gen3, stredná trieda

PCIe 4.0 ×4 / NVMe

Šírka pásma:	7 877 MB/s
Latencia:	~15 µs
Použitie:	Desktop, gaming, workstation

PCIe 5.0 ×4 / NVMe

Šírka pásma:	15 754 MB/s
Latencia:	~10 µs
Použitie:	Enterprise, AI servers (2024–)

CXL 3.0 (PCIe 5.0)

Šírka pásma:	~64 GB/s
Latencia:	<200 ns
Použitie:	Memory expansion, fabric (2025)

NVMe vs. AHCI – Architektúra protokolu a príkazové fronty

AHCI (Advanced Host Controller Interface)

- Navrhnuté pre mechanické HDD (2004)
- 1 príkazová fronta, max. 32 čakajúcich príkazov (NCQ)
- Komunikácia: CPU → AHCI reg → SATA PHY → disk
- Niekoľko registrov, polling/interrupt model
- PCIe AHCI (aka legacy NVMe): funguje aj na PCIe ale bez výhod NVMe
- Limituje výkon SSD: bottleneck na ~100k IOPS
- Stále bežný na starých platformách (Intel 100/200 series)

VS

NVMe (Non-Volatile Memory Express)

- Navrhnuté pre NAND Flash (spec. 1.0 = 2011, v2.0 = 2021)
- 65 535 front príkazov × 65 535 príkazov/fronte
- Priame PCIe pásmo: CPU ↔ NAND bez medzivrstiev
- Multi-core: každé jadro CPU má vlastnú I/O frontu
- NVMe 2.0: ZNS, FDP, Key-Value, Zoned Storage
- >1M IOPS v produkčných PCIe 5.0 jednotkách (2025)
- NVMe-oF (over Fabrics): RDMA/TCP rozšírenie do sietí

PCIe 5.0 NVMe v produkcii (2024–2025) a CXL 3.0

PCIe 5.0 NVMe – Stav (2025)

- Crucial T705 / WD Black SN8100 / Sabrent Rocket 5: 14.5 GB/s seq. čítanie, 1.8M IOPS
- Phison E26 kontrolér: 8-core ARM, 14 GT/s NAND interface, LDPC HW engine
- Tepelný limit: 18 W pod záťažou → aktívne chladenie nevyhnutné (bez heatsink: throttle na Gen3 rýchlosť)
- AMD Zen 5 + Intel Core 200S: natívna podpora PCIe 5.0 M.2
- Enterprise: attach rate PCIe 5.0 = >40 % nových serverov H1 2025
- Micron 6500 ION: 200+ layer NAND, 4.8 TB/in³ (5× hustejší než 24TB nearline HDD)

CXL 3.0 – Compute Express Link

- Protokol nad PCIe 5.0 pre koherentné CPU↔pamäť prepojenie
- CXL 1.1/2.0: CPU↔akcelerátor (GPU, FPGA) a memory expander
- CXL 3.0 (2022): fabric topológia, memory pooling naprieč servermi
- Latencia: <200 ns (vs. RDMA ~1–2 μs, NVMe ~10 μs)
- Samsung CXL DRAM moduly: 128–512 GB v jedinom CXL module (2024)
- Perspektíva: stiera hranice medzi pamäťou a úložiskom → Computational Storage

Fyzické formáty: M.2, U.2, E1.S / E3.S (EDSFF)

M.2 (NGFF)

Rozmer: 2230, 2242, 2260, 2280, 22110 mm. Konektor B (SATA/PCIe ×2), M (PCIe ×4/NVMe). Dominuje v notebookoch + desktopoch. Tepelný limit ~8–10 W bez heatsink.

U.2 (SFF-8639)

2.5" enterprise NVMe: hot-swap podpora, 25 W TDP. Formát pre data centre (Intel Optane P5800X, Samsung PM9A3). Postupne nahradzovaný E1.S v nových deploymentoch.

E1.S / E3.S (EDSFF)

Enterprise & Datacenter SSD Form Factor (SNIA/JEDEC). E1.S (15mm): server blade, 25 W. E3.S: vyššia kapacita. Meta Open Rack V3: mandátny E3.S. Alibaba 2025: 80 % procurement EDSFF.

Add-in Card (AIC)

PCIe karta plnej dĺžky: max výkon bez M.2 tepelných obmedzení. Seagate Nytro, Samsung PM9A3 AIC. 75 W TDP (PCIe slot power). Bežné v AI training serveroch.



Trhový posun k EDSFF (2025)

NVMe M.2 = >60 % dodávok PC SSD (2023). U.2 klesá v prospech E1.S/E3.S v dátových centrách. Microsoft Azure: E1.S redukuje chladiace náklady o 22 % vs. U.2. Meta Open Rack V3: mandátny E3.S standard.

Spôsoby záznamu a kódovania

RLL · PRML · ECC · Scrambler · ZBR · Write Amplification

07

Kódovanie dát na HDD – RLL, PRML a korekcia chýb

RLL (Run-Length Limited)

RLL 2,7 / 1,7

Obmedzuje dĺžku behov 0-bitov medzi po sebe idúcimi 1-bitmi. RLL (1,7): min. 1 nula, max. 7 núl za sebou → zlepšuje synchronizáciu clock recovery. RLL (2,7): štandard 80.–90. roky. Kodovacia účinnosť ~88–93 %.

PRML (Partial Response Maximum Likelihood)

Viterbi dekodér

Číta analógový signál s ISI (Inter-Symbol Interference) a rekonštruje bitovú sekvenciu pomocou Viterbiho algoritmu (Maximum Likelihood Sequence Detection). Štandard pre moderné HDD od ~1990. EPRML = rozšírený model odozvy kanála.

Scrambler (Whitening)

LFSR XOR

Pred zápisom XOR-uje dáta s pseudo-náhodnou sekvenciou (LFSR). Eliminuje dlhé behy rovnakých bitov → lepší frekvenčný spektrum signálu a synchronizácia. Povinný v Advanced Format (AF) diskoch.

ECC na HDD (Reed-Solomon → LDPC)

BCH → RS → LDPC

Historicky Reed-Solomon kódy (opravujú burst chyby). Moderné HDD: LDPC kódy (Low-Density Parity-Check) opravujú viacbitové chyby → umožňujú vyššie areal density pri zachovaní BER < 10^{-14} .

Zónové Bitové Záznamy (ZBR) a správa chýb – HDD vs. SSD

ZBR – Zone Bit Recording

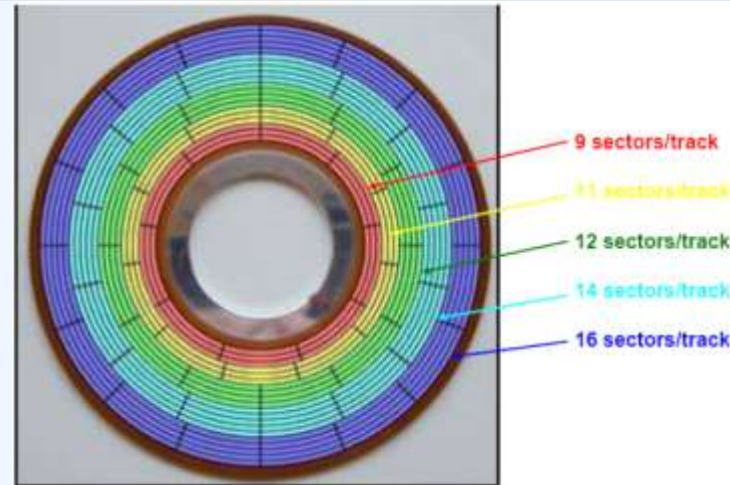
Vonkajšie stopy platne sú fyzicky dlhšie ako vnútorné. ZBR využíva túto skutočnosť: zapisuje viac sektorov na vonkajšie stopy. Efekt: čítacia rýchlosť na vonkajšej zóne je až 2× vyššia ako na vnútornej.

Správa chýb na HDD: S.M.A.R.T.

Self-Monitoring, Analysis and Reporting Technology. Monitoruje: seek error rate, reallocated sectors (slabé/chybné bloky), spin-up time, UDMA CRC, current pending sectors. Varovanie pred blížiacim sa zlyhaním disku.

NAND Error Management: Program Disturb & Read Disturb

Program Disturb: zápis do susedných buniek ovplyvňuje náboj v nezapisovaných bunkách. Read Disturb: opakované čítanie z rovnakej stránky mení stav susedných buniek. FTL počíta tieto udalosti a periodicky prepisuje postihnuté bloky.



NAND vs. HDD – Chybovosť

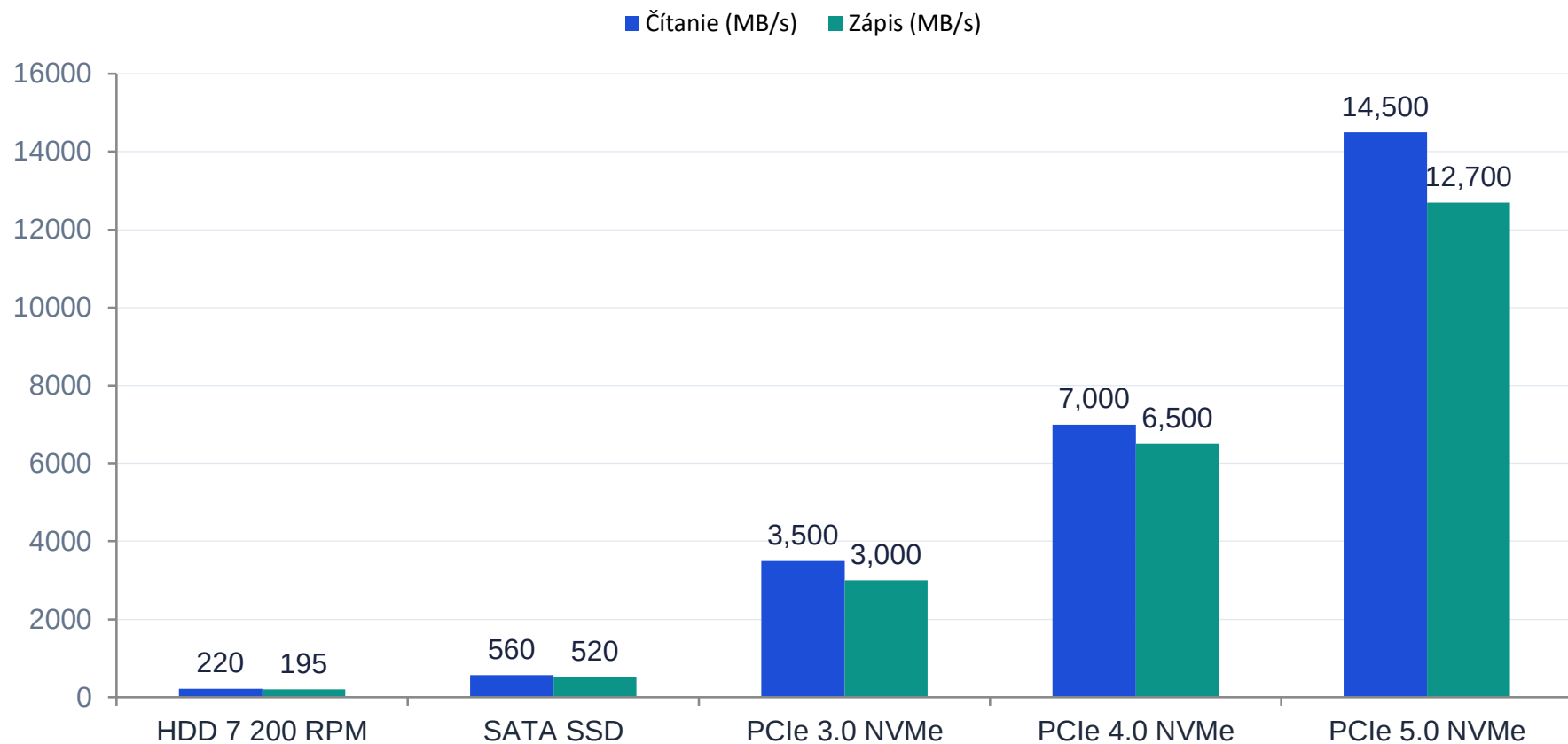
HDD raw BER	$\sim 10^{-8}$ (magnetický)
NAND TLC raw BER	$\sim 10^{-2}$ (pred ECC)
Po LDPC ECC	$< 10^{-14}$ (cieľ)
HDD MTBF	1–2 mil. hodín
SSD MTBF	~ 2 mil. hodín

Výkonnostné porovnanie & Trhové štatistiky

Benchmarky · IOPS · latencia · trhový podiel 2024–2025

08

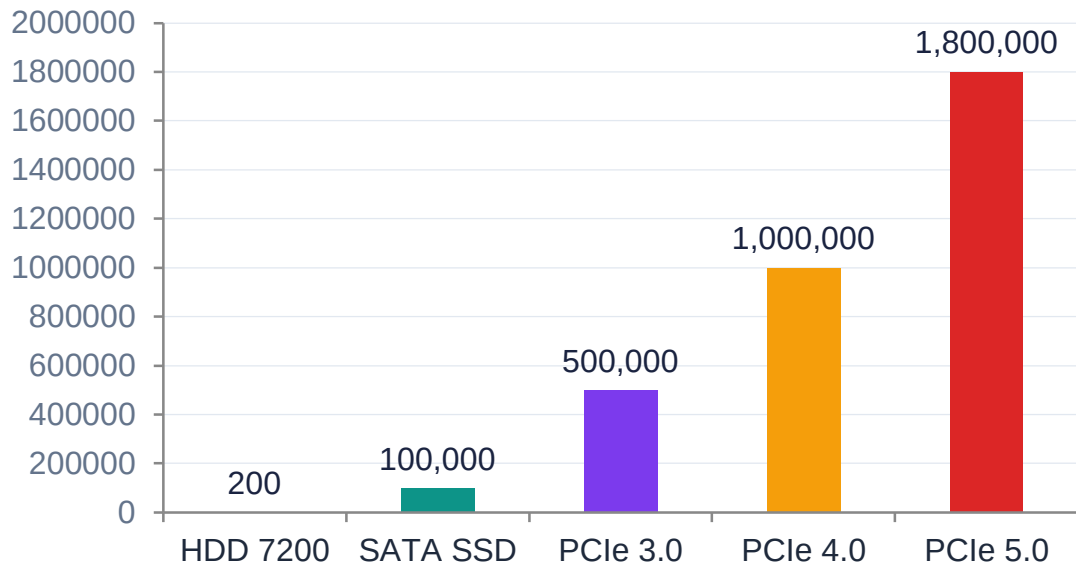
Výkonnostné porovnanie – Sekvenčná rýchlosť čítania a zápisu (MB/s)



Zdroj: Crucial T705 datasheet (2024), WD Black SN8100 spec, Samsung 870 EVO spec, Tom's Hardware PCIe 5.0 SSD reviews (2025) | Hodnoty pre najlepšie modely každej kategórie

Random 4K IOPS a latencia – Architektúrny dopad

Random 4K Read IOPS



Latencia (random 4K read): HDD ~5 ms · SATA SSD ~100 μs · PCIe 3.0 ~20 μs · PCIe 4.0 ~15 μs · PCIe 5.0 ~10 μs

Zdroj: Tom's Hardware PCIe 5.0 SSD Benchmarks (2025) | Crucial T705 datasheet | Western Digital WD Black SN8100 specs

1.8M

IOPS – PCIe 5.0 NVMe (Crucial T705, 2024)

200

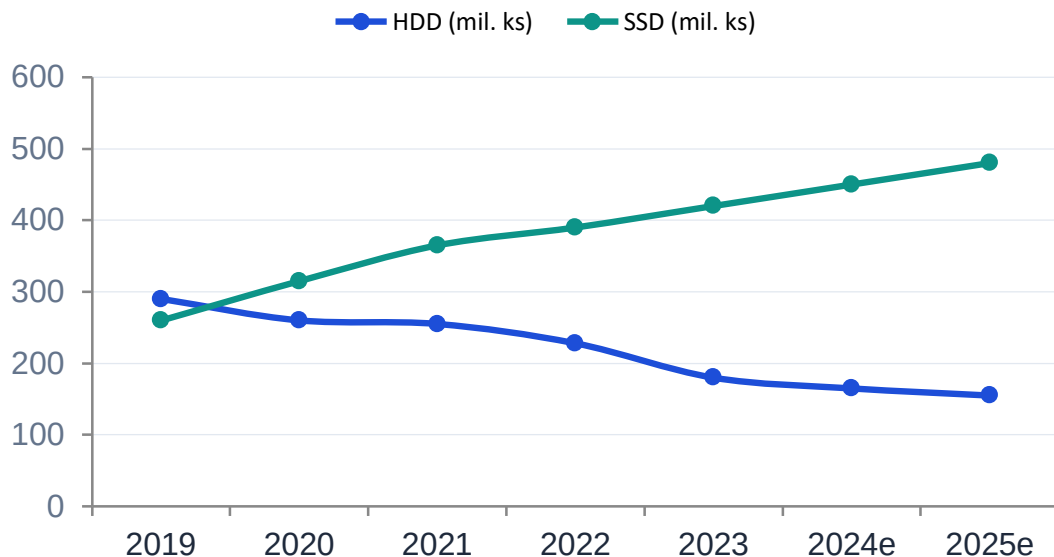
IOPS – HDD 7 200 RPM

×9 000

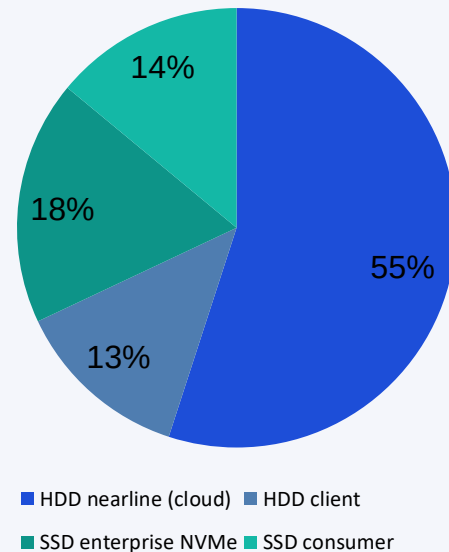
rozdiel v IOPS (HDD→PCIe5)

Trhové štatistiky úložísk – HDD vs. SSD (2019–2025)

Globálne dodávky (mil. kusov/rok)



Podiel kapacít 2025 (%)



Kľúčový záver: SSD prekročilo HDD v kusoch (~2022). HDD nearline dominuje v celkovej kapacite (archív, cloud cold tier). NVMe prekročilo 60 % podiel v enterprise arrayoch (2025). Seagate HAMR 30 TB: v sériovej výrobe 2024–2025. Trh SSD: \$31.3 mld. (2025) → \$76.4 mld. (2031) pri CAGR 16 %.

Súhrnné porovnanie HDD vs. SSD (2025) – architekúrna perspektíva

Vlastnosť	HDD (7 200 RPM)	SSD – SATA	SSD – NVMe PCIe 5.0	Výhoda
Seq. čítanie	80–250 MB/s	~550 MB/s	až 14 500 MB/s	NVMe x65
Latencia (rand.)	5–15 ms	~100 µs	~10 µs	NVMe x1 500
IOPS (4K rand.)	50–250	~100 000	až 1 800 000	NVMe x7 000+
Max. kapacita	32 TB (HAMR, 2025)	4 TB	8 TB (cons.), 122 TB (ent.)	HDD (nearline)
Cena/TB (2025)	~15–20 €	~55–80 €	~80–200 €	HDD 4–10×
Spotreba (idle/load)	2–5 W / 5–12 W	0.5 W / 2–4 W	0.5 W / 7–18 W	SATA SSD
Vibrácie & nárazy	Citlivý (HDD off)	Odolný	Odolný	SSD
Hlučnosť	20–35 dB	0 dB	0 dB	SSD
MTBF	1–2 mil. hod	2 mil. hod	2 mil. hod	Podobné

Budúcnosť ukladania – PCIe 6.0, HBF, DNA Storage, Computational Storage

PCIe 6.0 (2027+)

PAM4 modulácia, 64 GT/s, $\times 4 = 256$ GB/s.
Ratifikované jan 2022. Prvé produkty ~2027.
PCIe 7.0 vo vývoji (128 GT/s). Termálne výzvy:
signálna integrita pri PAM4.

HBF – High-Bandwidth Flash

SK Hynix + SanDisk (2025): NAND + DRAM
stacked (HBM štýl). Cieľ: 100M IOPS, sub- μ s
latencia. NVIDIA GTC 2025: storage I/O je
bottleneck AI pipeline.

DNA Data Storage

Microsoft Project Silica: sklenené tabuľky, 10
000+ rokov životnosť. DNA: 1 exabajt/mm³
(teoreticky). Čítanie/zápis: hodiny–dni → len
cold archív.

Computational Storage

Processing-in-storage: ARM/RISC-V jadrá
priamo v SSD. Redukuje data movement medzi
CPU a storage. Samsung SmartSSD, NGD
Catalina (2024). CXL + CS = budúci tier.

MRAM / PCM / ReRAM

Kandidáti na SCM (Storage Class Memory):
Everspin MRAM (sub-100 ns, 10^{15} P/E). Po
konci Optane (2022) hľadá priemysel nástupcu
medzi DRAM a NAND.

QLC → PLC NAND masovosť

QLC mainstream 2025 (Solidigm 122 TB,
Micron 6600 ION). PLC v laboratóriách.
Samsung V-NAND 1 000 vrstiev (cieľ 2030).
Hustota: 300+ layer = >4 Tb/Die.

Záver – 5 kľúčových poznatkov

1

Hierarchia pamätí = základ architektúry

Register → Cache → RAM → SSD → HDD: každá úroveň prináša iný kompromis rýchlosť/kapacita/cena. CPU bottleneck je storage latencia (ms vs. ns).

2

HDD = HAMR + kapacita + nízka cena/TB

HAMR Mozaic 3+ (30 TB, 2024) → Mozaic 4+ (40+ TB, 2025) → 100 TB/disk pred 2032. HDD 6× lacnejší/TB ako SSD. Nearline HDD = ~55 % kapacity v cloude.

3

SSD = 3D NAND + FTL + NVMe protokol

321-layer QLC (SK Hynix 2025), PCIe 5.0: 14.5 GB/s, 1.8M IOPS. FTL (WL + GC + ECC) skrýva NAND asymetriu. ZNS/FDP redukujú WAF na ~1.0.

4

Rozhrania určujú výkonnostný strop

SATA (600 MB/s) → NVMe PCIe 5.0 (14.5 GB/s) = 24× za ~10 rokov. CXL 3.0 stiera hranicu medzi pamäťou a úložiskom. NVMe >60 % enterprise arrayov (2025).

5

Kódovanie & záznam = základ spoľahlivosti

LDPC ECC, RLL/PRML, Wear Leveling, ZNS FDP – softvérovo-hardvérová kooperácia určuje skutočnú životnosť (TBW) a konzistentnosť latencie v produkčných systémoch.

Použitá literatura a zdroje

- [1] Patterson, D. A., Hennessy, J. L. – Computer Organization and Design (ARM Ed.), 5th ed., Morgan Kaufmann, 2016
- [2] Tanenbaum, A. S. – Modern Operating Systems, 4th ed., Pearson, 2014 – kap. 5: I/O
- [3] NVMe Base Specification 2.0 – NVMe Working Group, 2021 [nvmeexpress.org]
- [4] JEDEC JESD218B – SSD Requirements and Endurance Test Method, 2021
- [5] Seagate – Mozaic 3+ HAMR White Paper & "The nanomaterials enabling the AI era", 2024–2025 [seagate.com]
- [6] TechRadar – HAMR and SMR: 6.5TB/platter demo, Dec 2025 [techradar.com]
- [7] Tom's Hardware – Seagate HAMR Mozaic roadmap & PCIe Gen5 SSD reviews (2024–2025) [tomshardware.com]
- [8] TrendForce – SK Hynix 321-layer QLC mass production 2025; 400-layer roadmap, Aug 2024 [trendforce.com]
- [9] Mordor Intelligence – Solid State Drive Market Size 2025–2031, 2025 [mordorintelligence.com]
- [10] Fast Company – Is 2026 the year for SSD datacenter crossover? Dec 2025 [fastcompany.com]
- [11] Zhang et al. – ZNS: Avoiding the Block Interface Tax for Flash-based SSDs, USENIX ATC 2021
- [12] OrdinaryTech – Storage Wars 2025: PCIe Gen 5 NVMe, Dec 2025 [ordinarytech.ca]
- [13] Micron – 6500 ION 200+ Layer NAND Datasheet & Blog (2024) [micron.com]
- [14] CXL Consortium – CXL Specification 3.0, 2022 [computeexpresslink.org]
- 15 <https://www.youtube.com/watch?v=wtdnatmVdlg&t=332s>
- 16 specfortech.wordpress.com/hardware/storage-hardware

Ďakujem za pozornosť

Otázky?

Tobiáš Lehen | Počítačové Architektúry

Univerzita Cyrila a Metoda | 2025/2026

