

SVEUČILIŠTE U ZAGREBU
FILOZOFSKI FAKULTET
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE
ZNANOSTI

Ak. god. 2014./2015.

Dijana Medić

**Dugoročno očuvanje podataka na
fizičkoj razini zapisa**

Završni rad

Mentor: izv. prof. dr. sc. Hrvoje Stančić

Zagreb, rujan 2015.

SADRŽAJ:

POPIS SLIKA.....	3
1. UVOD	4
2. RAZINE DIGITALNIH PODATAKA.....	5
3. POHRANA PODATAKA.....	6
3.1. Hijerarhija memorije	7
3.2. Kategorizacija pohrane	12
3.2.1. Primarna pohrana	13
3.2.2. Sekundarna pohrana	14
3.2.3. Tercijarna pohrana.....	15
3.2.4. Neizravna pohrana.....	17
3.2.5. Izravna, poluizravna i neizravna pohrana.....	18
4. TVRDI DISK	19
4.1. Građa tvrdog diska	20
4.2. Pristup podacima na tvrdom disku.....	22
4.3. Lociranje podataka na tvrdom disku	22
4.4. Relativna učestalost pogrešaka i rukovanje	23
4.5. Redundantno polje nezavisnih diskova	25
4.5.1. RAID razine 0	26
4.5.2. RAID razine 1	27
4.5.3. RAID razine 2	28
4.5.4. RAID razine 3	29
4.5.5. RAID razine 4	30
4.5.6. RAID razine 5	31
4.5.7. RAID razine 6	32
4.5.8. RAID 0+1.....	33
4.5.9. RAID 1+0.....	34
6. PROPADANJE PODATAKA.....	35
6.1. Detektirano kvarenje podataka.....	37
6.2. Neotkriveno kvarenje podataka.....	38
6.2.1. Pogrešno upućen zapis	38
6.2.2. Rastrgan zapis	39
6.2.3. Oštećenje puta podataka.....	39
6.2.4. Onečišćenje pariteta	40
6.3. Protumjere.....	40
7. ZAKLJUČAK.....	42
8. LITERATURA	43
SAŽETAK.....	46
SUMMARY.....	47

POPIS SLIKA

Slika 1: Hijerarhija pohrane	9
Slika 2: Hijerarhija memorije	11
Slika 3: Podjela pohrane	13
Slika 4: Robotizirana biblioteka magnetskih traka	16
Slika 5: Građa tvrdog diska	21
Slika 6: RAID razine 0	26
Slika 7: RAID razine 1	27
Slika 8: RAID razine 2	28
Slika 9: RAID razine 3	29
Slika 10: RAID razine 4	30
Slika 11: RAID razine 5	31
Slika 12: RAID razine 6	32
Slika 13: RAID 0+1	33
Slika 14: RAID 1+0	34
Slika 15: Oštećena .jpeg datoteka u kojoj je većina podataka izgubljena	36

1. UVOD

U današnje vrijeme kada se govori o čuvanju digitalnih podataka svi pomalo imamo sindrom hrčka. Pohrana je jeftina, pa ako postoji makar mala šansa da bi nam neki podaci mogli biti korisni u budućnosti, mi ćemo ih zadržati. Znamo da mediji za pohranu nisu dovoljno pouzdani sami po sebi, stoga izrađujemo i čuvamo sigurnosne kopije. Ali što više podataka čuvamo i što ih duže čuvamo, veća je vjerojatnost da će neki od tih podataka biti neupotrebljivi kada ih zatrebamo. Iako ima cijeli niz razloga zašto s vremenom podaci mogu postati neupotrebljivi, ovaj rad će se zadržati samo na fizičkoj razini zapisa.

Cilj rada je predstaviti fizičku razinu zapisa kroz detaljni prikaz sustava za pohranu, razmotriti što je sustav za pohranu podataka, na koji način je posložen u arhitekturi računala, s naglaskom na tvrdi disk kao glavni medij sekundarne pohrane i na problemima koji se javljaju u očuvanju samog fizičkog zapisa u binarnom obliku. Bit će izneseni i tehnološki prijedlozi i rješenja u sprečavanju tih problema.

Rad je podijeljen u sedam cjelina koje zaokružuju temu. Nakon uvoda slijedi podjela digitalnih podataka na tri razine, kako bi se upoznali sa samom fizičkom razinom podataka.

Zatim slijedi poglavlje o pohrani podataka, te razlika između pohrane i memorije. Hijerarhija memorije pruža bolje razumijevanje u pristupu uređajima pohrane, nakon čega dolazi podjela pohrane, čime se omogućuje pristup temi tvrdog diska kao glavnom mediju u ovom radu. Nakon što je iznijeta građa tvrdog diska i pristup podacima na njemu, uvode se neki problemi očuvanja podataka na fizičkoj razini u obliku pogrešaka diskova i rukovanja njima.

Nakon samog tvrdog diska prolazi se cijeli niz razina redundantnih polja nezavisnih diskova, upoznavanje sa svakom razinom i prednostima koje one nude u očuvanju podataka.

U posljednjem poglavlju se navodi problematika kvarenja podataka i propadanja bitova čak i kada nisu u upotrebi, različite vrste kvarenja podataka, te neke od mjera zaštite od tih kvarenja.

2. RAZINE DIGITALNIH PODATAKA

Svi digitalni objekti imaju svoja osnovna svojstva bez obzira o kakvom se tipu digitalnog podatka radi. Na temelju tih osnovnih svojstava, svaki digitalni objekt se može promatrati na tri razine. To su fizička, logička i konceptualna razina.¹

Na fizičkoj razini digitalni podatak je jednostavno zapis znakova na određeni medij pohrane u obliku binarnog sustava, tj. u obliku niza jedinica i nula. Svaki podatak se digitalno bilježi binarnim sustavom znakova, ali ovisno o fizičkim svojstvima samog medija pohrane, način na koji se digitalni podatak bilježi, razlikuje se od medija do medija. Osim što se način zapisa razlikuje između različitih medija na koje se pohranjuje, razlikuje se i način pohrane unutar iste vrste medija.²

Na logičkoj razini, objekt postaje prepoznatljiv softveru i može se procesuirati. Logička razina određuje način na koji se sadržaj fizički organizira i zapisuje, dok vrstu medija i zapis na taj medij zanemaruje.

Konceptualna razina je razina na kojoj objekt dobiva smisao, te se prepoznaje kao informacijska jedinica - npr. dokument, knjiga, slika, pjesma, film...

Sve tri razine su međusobno povezane i ovise jedna o drugoj. Fizička razina je sam zapis na medij pohrane, dok logička određuje način na koji se podatak zapisuje. Logička razina mora također biti fizički zabilježena. Sadržaj i struktura konceptualne razine moraju biti sadržani u logičkoj razini. Konceptualna razina može biti organizirana na različite načine, o čemu ovisi i njezina interpretacija.

Iako su sve razine usko povezane i svaka od njih je od jednake važnosti kad je u pitanju očuvanje digitalnog objekta, ovaj rad će se zadržati isključivo na fizičkoj razini samog zapisa i iznijeti probleme koji se javljaju tamo i kako se boriti s tim problemima.

¹ Thibodeau, K. Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. *The State of Digital Preservation: An International Perspective*, Council on Library and Information Resources (CLIR), Washington, D.C., SAD, srpanj 2002.; str. 4-15.

² Stančić, H. Arhivsko gradivo u elektroničkom obliku: mogućnosti zaštite i očuvanja na dulji vremenski rok. *Arhivski vjesnik*, No. 49, prosinac 2006.; URL: http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=9508. (12.8.2015.)

3. POHRANA PODATAKA

Pohrana računalnih podataka (engl. *computer data storage*), poznata i pod nazivima kao pohrana (engl. *storage*) i kao memorija (engl. *memory*), je tehnologija koja se sastoji od uređaja za pohranu podataka i medija za pohranu podataka, te se koriste za čuvanje digitalnih podataka. To je jezgrena funkcija i temeljna komponenta računala.³

Iako se pohrana ponekad naziva i memorija trebalo bi ih razlikovati. Najčešće se brze, nepostojane tehnologije (koje gube podatke kada se isključe iz struje) nazivaju memorija, dok se sporije, postojeane tehnologije nazivaju pohrana. Na žalost, kada se govori o trajnoj pohrani, često se upotrebljava termin memorija. Razlika između pohrane i memorije je ta da se postojana (engl. *nonvolatile*) pohrana koristi za čuvanje programa i podataka sve do njihove namjerne promjene ili brisanja od strane korisnika, dok je nepostojana (engl. *volatile*) memorija privremeni radni prostor za dohvaćanje programa, izvršavanje instrukcija i obradu podataka.

Pohrana implicira repozitorij koji zadržava svoj sadržaj i nakon što se prekine dovod struje, te je time postojano mjesto gdje se čuvaju digitalni podaci dok se namjerno ne obrišu. Pohrana uglavnom podrazumijeva magnetske diskove, SSD (*Solid State Drive*) diskove i USB (*Universal Serial Bus*) uređaje, no termin se koristi i za magnetske trake i optičke diskove.⁴

Memorija, koja se sastoji od DRAM (*Dynamic-Random Access Memory*) i SRAM (*Static-Random Access Memory*) čipova je privremeni radni prostor za izvršavanje uputa i obradu podataka i ti čipovi ne pamte sadržaj nakon što se napajanje isključi. Kao glavni resurs u računalu, memorija određuje veličinu i broj programa koji se mogu izvoditi u isto vrijeme, kao i količinu podataka koji se mogu trenutno obrađivati. Sve obrade podataka i izvršavanja programa se odvijaju u memoriji na način da se upute programa kopiraju u memoriju s diska ili mreže, zatim se izdvajaju u kontrolni sklop procesora na analizu i izvršenje. Upute zatim usmjeruju računalu na unos, obradu i prikaz podataka. Kako se podaci unose u memoriju, prethodni sadržaj tog prostora se gubi. Nakon što su podaci u memoriji, mogu se obrađivati, a rezultat se šalje na ekran, printer, u pohranu ili na mrežu.⁵

³ Computer Data Storage. URL: https://en.wikipedia.org/wiki/Computer_data_storage. (13.8.2015.)

⁴ Storage. URL: <http://www.yourdictionary.com/storage>. (14.8.2015.)

⁵ Memory. URL: <http://www.yourdictionary.com/memory>. (14.8.2015.)

Memorija je važan resurs i stoga se ne smije uzalud trošiti. Mora biti dodijeljena od strane operativnog sustava kao i od strane aplikacija, a zatim otpuštena kada više nije potrebna. Zalutali programi mogu zauzeti memoriju i ne puštati je, što rezultira sa sve manje i manje memorije dostupne drugim programima. Ponovno pokretanje računalne memorije daje svjež početak, zbog čega ponovno podizanje sustava rješava mnoge probleme s aplikacijama. Osim toga, ako operativni sustav ima bugova (grešaka u kôdu programa), neispravna aplikacija se može upisivati u dio memorije koju već koristi neki drugi program, uzrokujući nespecificirano ponašanje poput zamrzavanja cijelog sustava.

3.1. Hijerarhija memorije

Hijerarhija memorije je raspon memorije i pohrane u računalnom sustavu.

Memorija je esencijalna za rad računalnog sustava, a ništa nije važnije za razvoj modernog memorijskog sustava od koncepta memorijske hijerarhije. Linearni memorijski sustav može biti zanimljiv zbog svoje jednostavnosti. Ali dobro implementirana hijerarhija omogućuje memorijskom sustavu da istovremeno pristupa: performansama najbržih komponenata, cijeni po bitu najjeftinijih komponenata i potrošnji energije energetski najučinkovitijih komponenata.

Korištenje memorijske hijerarhije je vrlo pogodno, u smislu da pojednostavljuje proces projektiranja memorijskih sustava. Upotreba hijerarhije dopušta dizajnerima da tretiraju dizajn sustava kao modularni proces, tj. da tretiraju memorijski sustav kao apstrakciju i na taj način mogu optimizirati pojedine podsustave (predmemoriju, dinamički RAM, diskove) odvojeno.

Moderna hijerarhija se sastoji od sljedećih komponenti od kojih svaka obavlja određenu funkciju u sustavu: predmemorije (engl. *cache*) koja uključuje SRAM, zatim od DRAM-a, i na kraju, diska.

Predmemorija (SRAM) pruža pristup naredbama programa i podacima koji imaju vrlo nisku latenciju (npr. četvrtinu nanosekunde po pristupu) i vrlo veliku širinu propusnog pojasa (engl. *bandwidth*) (npr. blok naredbi od 16 bajtova i blok podataka veličine 16 bajtova po ciklusu, što onda iznosi 32 bajta po četvrtini nanosekunde, odnosno 128 bajtova po

nanosekundi, tj. 128 GB/s). Također je važno napomenuti da *cache* ima relativno niske energetske potrebe u odnosu na druge tehnologije.

DRAM (dinamička memorija nasumičnog pristupa) pruža pohranu nasumičnog pristupa koja je relativno velika, relativno brza i relativno jeftina. Velika je i jeftina u odnosu na *cache* i brza u odnosu na disk. Njena glavna značajka je ta da je dovoljno brza i dovoljno jeftina da služi kao radna memorija.

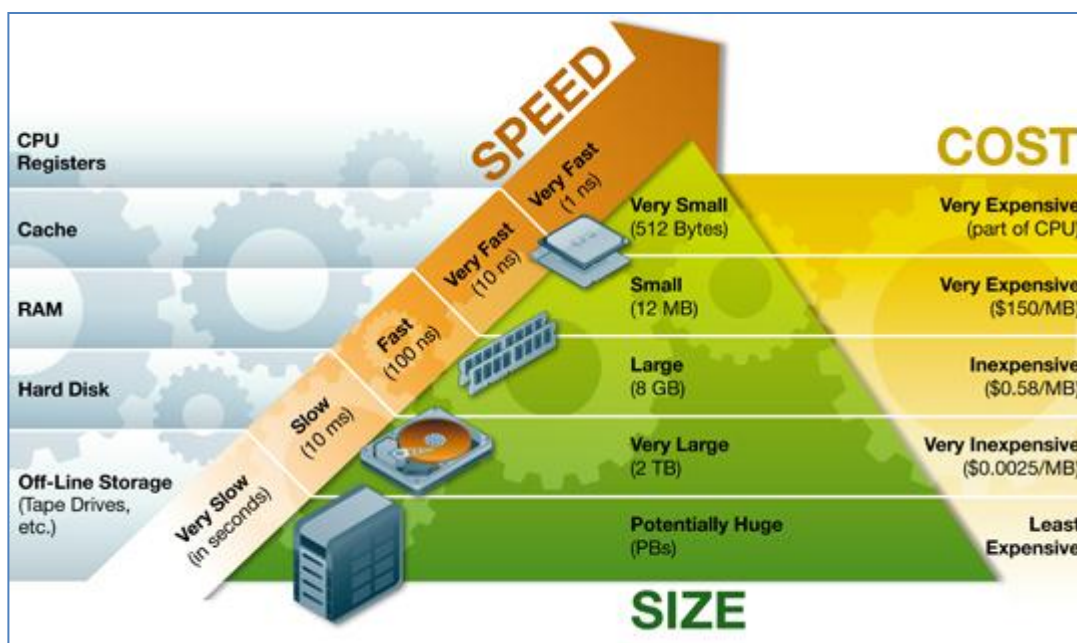
Disk omogućuje trajnu pohranu po izrazito niskim cijenama po bitu. Gotovo svi računalni sustavi očekuju da većina podataka bude trajna, ali i da se daju mijenjati, tako da memorijski sustav na nekoj razini mora imati trajnu pohranu. Prednost diska je njegova vrlo razumna cijena, koja je dovoljno niska da korisnici mogu kupiti dovoljno mjesta za pohranu na tisuće pjesama, video isječaka, slika i drugih digitalnih objekata koje žele akumulirati i imati u svom posjedu.

Zašto od navedenih tehnologija nije moguće napraviti linearan, jednostavan sustav memorije? Predmemorija je daleko preskupa tehnologija da bi se koristila za trajnu pohranu. Toliko je skupa da bi trošak pohrane audio zapisa jednog albuma pjesama premašio cijenu originalnog CD-a za nekoliko redova veličine. Disk je, s druge strane, daleko prespor da se koristi kao radna memorija jer se njegovo vrijeme traženja nasumičnih pristupa mjeri u milisekundama. Od te tri tehnologije, DRAM je najbliže da se koristi sâm kao sustav memorije. Dovoljno je brz da i bez potpore predmemorije može djelovati kao radna memorija za mnoge ugrađene sustave (engl. *embedded systems*), a uz potporu baterije može funkcionirati i kao trajna pohrana. Međutim, DRAM sâm nije dovoljno jeftin za potrebe korisnika, koji često žele terabajte trajne pohrane, a i s vremenom nasumičnog pristupa od desetak nanosekundi, DRAM nije dovoljno brz da služi kao jedina memorija za moderne sveobuhvatne mikroprocesore, koji preferiraju novi blok instrukcija svaki djelić nanosekunde.

Za sada nema tehnologije koja pruža sve željene karakteristike: nisku cijenu, postojanost, veliku širinu propusnog pojasa, nisku latenciju itd. Upravo iz tog razloga gradimo sustav u kojem je svaka komponenta dizajnirana da nudi jednu ili više karakteristika, te upravljamo radom sustava tako da slabije karakteristike svake tehnologije ostanu "skrivena". Na primjer, ako većinom memorijskih referenci koje poziva mikroprocesor, barataju predmemorijski i/ili DRAM podsustavi, onda će se disk koristiti rijetko, i stoga, njegova ekstremno velika latencija će vrlo malo pridonijeti prosječnom vremenu pristupa. Ako se većina podataka nalazi u diskovnom podsustavu, a vrlo malo ih je u upotrebi u bilo kojem

danom trenutku, tada predmemorijski i DRAM podsustavi neće trebati mnogo prostora pohrane, te će njihova cijena po bitu vrlo malo utjecati na prosječnu cijenu cijelog sustava. Ako je dobro složena hijerarhija, memorijski sustav će imati cijenu koja se približava cijeni najdonjeg sloja, te prosječno vrijeme pristupa i širinu propusnog pojasa koja se približava onoj najgornjeg sloja.⁶

Gotovo sva računala u praksi koriste hijerarhiju pohrane koja stavlja brze, ali male i skupe opcije za pohranu bliže procesoru (engl. *central processing unit* - CPU), a sporije ali veće i jeftinije opcije, dalje od procesora⁷. (slika 1)



Slika 1: Hijerarhija pohrane

Preuzeto sa: http://www.ts.avnet.com/uk/products_and_solutions/storage/hierarchy.html
(13.8.2015.)

Slika 1 prikazuje kako su brže memorije bliže procesoru, a sporije i jeftinije memorije ispod njih i dalje od procesora. Cilj je predstaviti korisniku što je više moguće memorije u najjeftinijoj tehnologiji, pružajući pristup brzinom koju nudi najbrža memorija. Podaci su

⁶ Jacob, B.; Ng, S. W.; Wang, D.T.; *Memory Systems: Cache, DRAM, Disk*. Elsevier Inc., Burlington, Massachusetts, USA, 2008.

⁷ Storage hierarchy. URL: http://www.ts.avnet.com/uk/products_and_solutions/storage/hierarchy.html.
(13.8.2015.)

slično hijerarhijski posloženi: razina bliže procesoru je generalno podskup bilo koje razine dalje, a svi podaci su pohranjeni na najnižoj razini. Nadalje, što se više udaljavamo od procesora, pristup razinama traje progresivno duže.⁸

Kako se spuštamo po hijerarhiji dešava se sljedeće:

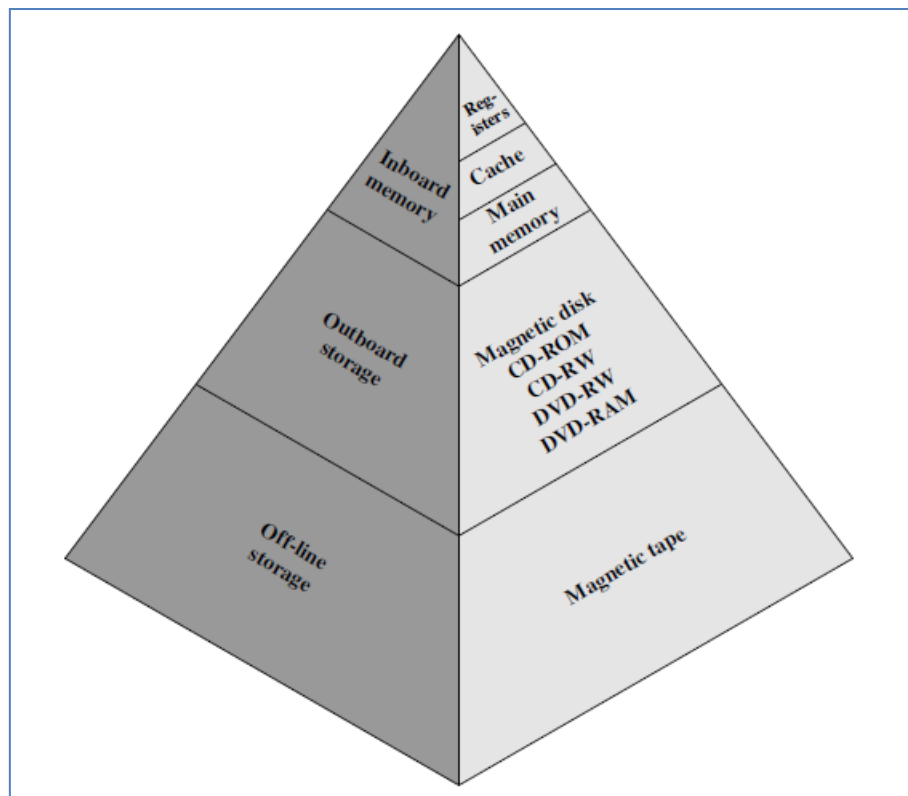
- a) smanjenje cijene po bitu,
- b) povećanje kapaciteta,
- c) povećanje vremena pristupa,
- d) smanjenje učestalosti pristupa memoriji od strane procesora.

Dakle, manje i skuplje memorije se nadopunjavaju većim, jeftinijim i sporijim memorijama odnosno pohranama. Ključ za uspjeh ove organizacije je točka d): smanjenje učestalosti pristupa. Primjenjujući razne tehnologije postoji cijeli spektar memorijskih sustava koji zadovoljava uvjete a) do d). Načelo poznato kao lokalnost reference (engl. *locality of reference*) je temelj da bi uvjet d) vrijedio. Tijekom izvođenja nekog programa, memorijske reference imaju tendenciju da se sakupljaju u klastere (engl. *clusters*). Programi obično sadrže niz ponavljajućih petlji i potprograma. Jednom kad se uđe u petlju ili potprogram, javljaju se ponavljajuće reference na mali skup uputa. Slično tome, operacije na tablicama i poljima (engl. *array*) uključuju pristup grupiranim skupovima podataka. Kroz duži period vremena klasteri koji su u upotrebi se promijene, ali u kratkom vremenskom razdoblju, procesor prvenstveno radi s fiksnim klasterima memorijskih referenci.

S tim u skladu, moguće je organizirati podatke pomoću hijerarhije, tako da je postotak pristupa svakoj uzastopce nižoj razini znatno manji od postotka pristupa razini iznad. Uzmimo za primjer memorijsku hijerarhiju s dvije razine. Neka memorija na drugoj razini sadrži sve naredbe programa i sve podatke. Trenutni klasteri mogu privremeno biti smješteni na prvoj razini. S vremena na vrijeme, jedan od klastera na prvoj razini će morati biti prebačen natrag na drugu razinu, kako bi se napravilo mjesta za novi klaster koji dolazi na prvu razinu. Međutim, u prosjeku će većina referenci biti na naredbe i podatke sadržane na prvoj razini.⁹

⁸ Patterson, D. A.; Hennessy, J. L. *Computer Organization and Design: The Hardware/Software Interface*. Fifth Edition. Elsevier Inc., Waltham, Massachusetts, USA, 2014.; str. 372-379.

⁹ Stallings, W. *Computer Organization and Architecture: Designing for Performance*. Eight Edition. Pearson Prentice Hall, Upper Saddle River, New Jersey, USA, 2010.; str. 111-118.



Slika 2: Hijerarhija memorije

Preuzeto iz: Stallings, W. Computer Organization and Architecture: Designing for Performance. Eight Edition. Pearson Prentice Hall, Upper Saddle River, New Jersey, USA, 2010.; str. 115.

Ovaj princip se može primijeniti i na više od dva nivoa memorije kao što je sugerirano hijerarhijom prikazanom na slici 2. Najbrži, najmanji i najskuplji tip memorije sastoji se od registara unutar procesora. Tipično procesor sadržava nekoliko desetaka takvih registara, iako neki strojevi sadrže i stotine registara.

Dvije razine niže nalazi se glavna memorija (engl. *main memory*) koja je glavni unutarnji memorijski sustav računala. Svaka lokacija u glavnoj memoriji ima jedinstvenu adresu.

Glavna memorija je obično proširena bržom i manjom predmemorijom. To je brza memorija jako malog kapaciteta (od 4 KB-a do par MB-a) i smještena je na CPU. Ona gotovo izravno podržava brzinu rada procesora i time daje privid da cijela glavna memorija radi brzinom predmemorije. To je uređaj za određivanje kretanja podataka između glavne

memorije i registara procesora, koji služi poboljšanju učinkovitosti. Tri oblika memorije, koji su upravo opisani, su nepostojeane memorije i baziraju se na poluvodičkim tehnologijama.

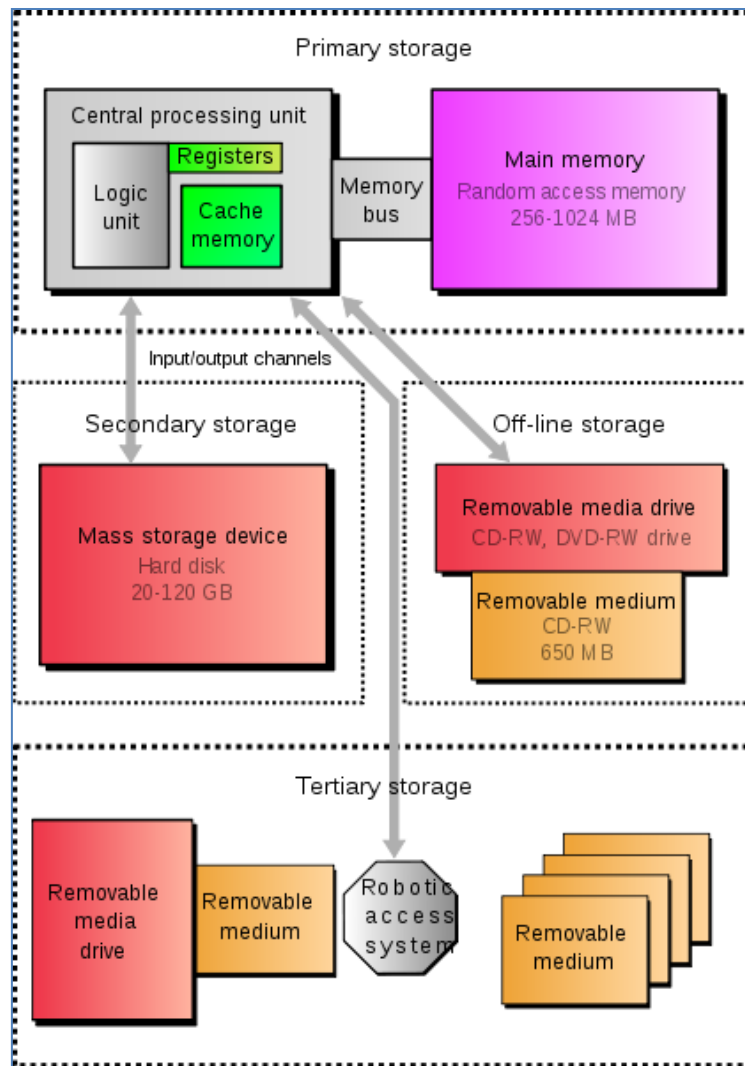
Podaci se trajnije pohranjuju na vanjske uređaje za pohranu, od kojih su najčešći tvrdis disk, te prijenosni mediji poput prijenosnog magnetskog diska, magnetskih traka ili optičkih medija. Vanjska, postojana memorija se još naziva i sekundarna memorija ili pomoćna memorija. Ona se koristi za pohranu programa i podatkovnih datoteka i obično ih vidimo u smislu datoteka i zapisa, a ne kao individualne bajtove. Disk se također koristi za proširenje glavne memorije, što nazivamo virtualna memorija (engl. *virtual memory*).

U hijerarhiju se također mogu uvrstiti i drugi oblici memorije. Na primjer, velika IBM-ova glavna računala (engl. *mainframes*) uključuju oblik unutarnje memorije poznat pod nazivom proširena pohrana (engl. *expanded storage*). Proširena pohrana koristi tehnologiju poluvodiča koja je sporija i jeftinija od one koju koristi glavna memorija. Strogo govoreći, ova memorija ne spada izravno u hijerarhiju, nego se promatra kao bočna grana: podaci se mogu prebacivati između glavne memorije i proširene pohrane, ali ne i između proširene pohrane i vanjske memorije.

Na posljetku, softverski se u hijerarhiju mogu dodati i dodatne razine. Dio glavne memorije se može koristiti kao međuspremnik za privremeno držanje podataka koji se trebaju učitati na disk. Takva tehnika se naziva memorija za ubrzanje pristupa disku (engl. *disk cache memory*) i poboljšava performanse na dva načina. Zapisi na disk se grupiraju, pa umjesto mnogo malih transfera podataka, imamo nekoliko velikih transfera podataka, što poboljšava performanse diska i smanjuje angažman procesora. Neki podaci namijenjeni da budu zapisani na disk, mogu biti pozvani od strane programa prije sljedećeg ispisa iz memorije na disk. U tom slučaju, podaci se učitavaju brzo iz memorije za ubrzanje pristupa disku, umjesto da se učitavaju sporo s diska.

3.2. Kategorizacija pohrane

Za bolje razumijevanje pohrane, ona se može kategorizirati kao: primarna, sekundarna, tercijarna i neizravna pohrana. Primarna i sekundarna pohrana su izravna pohrana, dok je tercijarna pohrana poluizravna.



Slika 3: Podjela pohrane

Preuzeto sa:

https://upload.wikimedia.org/wikipedia/commons/3/3e/Computer_storage_types.svg
(18.8.2015.)

3.2.1. Primarna pohrana

Primarna pohrana (engl. *primary storage*) je najviša razina i sastoji se od procesorskih registara, predmemorije i memorije. To su jedine komponente koje su izravno dostupne središnjoj procesorskoj jedinici (CPU). CPU može kontinuirano čitati podatke pohranjene u njima i izvršavati sve naredbe na brz i ujednačen način. Još se nazivaju i unutarnja memorija,

glavna memorija ili primarna memorija. Primarni uređaj za pohranu je medij koji drži podatke kratko vrijeme dok je računalo uključeno.

Kao što je prikazano na slici 3, postoje još dva podsloja primarne pohrane uz glavni RAM velikog kapaciteta. Procesorski registri se nalaze unutar procesora. Svaki registar obično sadrži 32 ili 64 bita. CPU nalaže aritmetičkoj i logičkoj jedinici (ALU) upute za obavljanje raznih izračuna ili drugih operacija. Registri su najbrži od svih oblika za pohranu računalnih podataka. Procesorska predmemorija je posrednik između izrazito brzih registara i dosta sporije glavne memorije. Uvedena je isključivo za poboljšanje performansi računala. Većina aktivno korištenih informacija u glavnoj memoriji se samo duplicira u predmemoriju koja je brža, ali znatno manjeg kapaciteta. S druge strane, glavna memorija je znatno sporija, ali ima puno veći kapacitet za pohranu od procesorskih registara.

Glavna memorija je izravno ili neizravno spojena na središnju procesorsku jedinicu preko memorijske sabirnice (engl. *memory bus*). Zapravo postoje dvije sabirnice: adresna sabirnica i podatkovna sabirnica. CPU prvo šalje memorijsku adresu koja ukazuje na željenu lokaciju podataka putem adresne sabirnice. Potom čita ili zapisuje podatke u memorijske ćelije, koristeći podatkovnu sabirnicu. Dodatno postoji i jedinica za upravljanje memorijom (engl. *memory management unit - MMU*). To je mali uređaj između CPU-a i RAM-a koji preračunava stvarnu memorijsku adresu.

S obzirom da je RAM nepostojana memorija (briše se kada se računalo isključi), računalo koje bi imalo samo takav tip memorije ne bi imalo izvor za čitanje uputa koje pokreću računalo. Stoga se za dizanje računala koristi postojana primarna pohrana koja sadrži mali program za podizanje (BIOS – *Basic Input/Output System*). BIOS se koristi da učita veći program iz postojane sekundarne pohrane u RAM, te da ga počne izvoditi. Postojana memorija koja se koristi za tu svrhu se naziva ROM – Read-Only Memory (memorija iz koje se podaci mogu samo čitati).

3.2.2. Sekundarna pohrana

Sekundarna pohrana (engl. *secondary storage*) se razlikuje od primarne time što nije izravno dostupna procesoru. Sustav koristi ulazno/izlazne – U/I (engl. *input/output – I/O*)

kanale koji, po potrebi i na zahtjev, kontroliraju protok podataka kroz sustav za spajanje na sekundarnu pohranu.

Sekundarna pohrana je postojana i ne gubi podatke kada se isključi iz struje, pa time i moderni računalni sustavi imaju tendenciju da imaju više sekundarne nego primarne pohrane. Sekundarna pohrana se još naziva i pomoćna memorija, pričuvna pohrana ili vanjska memorija. S obzirom da je vanjski uređaj i ne pristupa joj se preko ulazno-izlaznih kanala nije izravno dostupna središnjoj procesorskoj jedinici. Skoro svu sekundarnu memoriju danas sačinjavaju tvrdi diskovi (engl. *hard disk drive* - *HDD*), koji su često spojeni u konfiguracije redundantnih polja neovisnih diskova (engl. *redundant array of independent disks* - *RAID*). Neki drugi primjeri sekundarne pohrane su brze fleš memorije (engl. *flash memory*), npr. univerzalna serijska sabirnica (engl. *Universal Serial Bus* – *USB*). Prije su se, kao uređaji vanjske memorije, koristile i diskete (engl. *floppy disk drive* – *FDD*), te ZIP diskete.

3.2.3. Tercijarna pohrana

Tercijarna pohrana (engl. *tertiary storage*) se uglavnom koristi za sigurnosne kopije podataka (engl. *backup*) i arhiviranje podataka. Iako se temelji na najsporijim uređajima, može se svrstati kao najvažnija pohrana u smislu zaštite podataka od raznih nepogoda koje mogu utjecati na informatičku infrastrukturu.

Većina uređaja u tercijarnoj pohrani je automatizirana preko robotike i softvera, kako bi se smanjili troškovi upravljanja i rizici ljudske pogreške. Robotski mehanizam postavlja i skida prijenosne medije za masovnu pohranu u skladu sa zahtjevima sustava. Često se ti podaci kopiraju na sekundarnu pohranu prije uporabe.

Tercijarna pohrana se prvenstveno koristi za arhiviranje podataka koji se rijetko koriste, budući da je puno sporija od sekundarne pohrane. To je najkorisnije za izuzetno velike pohrane podataka kojima se pristupa bez ljudskih subjekata. Tipični primjer su biblioteke magnetskih ili optičkih traka (engl. *tape libraries*) i diskovni automati (engl. *jukeboxes*).

Kada računalo treba pročitati podatke iz tercijarne pohrane, prvo konzultira katalošku bazu podataka da utvrdi koji disk ili traka sadrži podatke koje treba. Nadalje, računalo daje naredbe robotskoj ruci da dohvati medij i stavi ga u pogon (engl. *drive*). Kada je računalo

gotovo sa iščitavanjem podataka, robotska ruka vraća medij natrag na njegovo mjesto u biblioteci.¹⁰ Tercijarna pohrana je također poznata i kao poluizravna (engl. *near-line*) pohrana.



Slika 4: Robotizirana biblioteka magnetskih traka

Preuzeto sa: <http://nsidc.org/monthlyhighlights/2009/12/nsidc-daac-data-pick-up-speed/>
(21.8.2015.)

¹⁰ Computer Data Storage. URL: https://en.wikipedia.org/wiki/Computer_data_storage. (20.8.2015.)

3.2.4. Neizravna pohrana

Neizravna pohrana (engl. *off-line storage*) je pohrana računalnih podataka na mediju ili uređaju koji nije pod kontrolom procesorske jedinice.¹¹ To je posljednja kategorija i tu spadaju prijenosne kasetne magnetske trake (engl. *removable tape cartridges*) i optički mediji poput CD-a, DVD-a, i BluRay-a. Neizravna pohrana se koristi za prijenos podataka između sustava, ali također omogućuje da podaci budu na sigurnom izvan sustava, te zahvaljujući tome tvrtke uvijek mogu imati osiguranu i spremljenu kopiju vrijednih podataka u slučaju katastrofe.

Neizravna pohrana funkcionira tako da je medij snimljen najčešće u sekundarnom ili tercijarnom uređaju za pohranu, a zatim fizički uklonjen ili iskopčan. Mora biti umetnut ili spojen od strane ljudskog operatera prije nego mu računalno može ponovo pristupiti. Za razliku od tercijarne pohrane, neizravnoj pohrani se ne može pristupiti bez ljudske interakcije.

Najviše se koristi za prijenos podataka, s obzirom da se samostalan medij može lako fizički transportirati. Kao što je već napomenuto, u slučaju da katastrofa, primjerice požar, uništi izvorne podatke, medij na udaljenoj lokaciji će po svojoj prilici ostati netaknut, te će se tako omogućiti oporavak od katastrofe. Neizravna pohrana povećava opću sigurnost informacija, budući da je fizički nedostupna računalu, pa tajnost i integritet podataka ne mogu biti oštećeni tehnikama napada na računalnoj bazi. Također, ako se informacijama pohranjenim u svrhu arhiviranja rijetko pristupa, neizravna pohrana je jeftinija od tercijarne pohrane.

U suvremenim osobnim računalima, većina sekundarnih i tercijarnih medija se koriste i za neizravnu pohranu. Optički diskovi i uređaji flash memorije su najpopularniji, a u nešto manjoj mjeri i prijenosni tvrdi diskovi. Stariji primjeri su diskete, Zip diskete i bušene papirnate kartice (engl. *punched cards*).¹²

¹¹Federal Standard 1037C – Telecommunications: Glossary of Telecommunication Terms. URL: <http://www.its.bldrdoc.gov/fs-1037/fs-1037c.htm> (21.8.2015.)

¹² Computer Data Storage. URL: https://en.wikipedia.org/wiki/Computer_data_storage. (21.8.2015.)

3.2.5. Izravna, poluizravna i neizravna pohrana

Formalna razlika između izravne, poluizravne i neizravne pohrane:

- izravna (engl. *on-line*) pohrana je odmah dostupna za U/I. Uključuje DRAM memorije i SSD diskove. Primarna i sekundarna pohrana su obje izravna pohrana.
- poluizravna (engl. *near-line*) pohrana nije odmah dostupna, ali se može učiniti izravnom bez ljudske intervencije. Uključuje optičke diskovne automate, automatizirane biblioteke magnetskih ili optičkih traka, te masivna polja diskova u mirovanju (engl. *massive array of idle disks – MAID*). Tercijarna pohrana je poluizravna.
- neizravna (engl. *off-line*) pohrana nije odmah dostupna i zahtijeva ljudsku intervenciju kako bi postala dostupna tj. izravna. To uključuje USB memorijske uređaje, optičke medije (*Compact Disk – CD, Digital Versatile Disk – DVD i BluRay Disk - BD*), prijenosne trake i ostale prijenosne medije¹³

Na primjer, tvrdi diskovi su izravna pohrana, dok su diskovi u MAID konfiguraciji poluizravna pohrana. Prijenosni mediji, kao što su kasetne magnetske ili optičke trake u bibliotekama magnetskih ili optičkih traka koje se mogu automatski učitati, su primjer poluizravne pohrane, dok su kasetne magnetske ili optičke trake koje se moraju ručno učitati, primjer neizravne pohrane.

¹³ Pearson, T.; *The Correct Use of the term Nearline*. URL: https://www.ibm.com/developerworks/community/blogs/InsideSystemStorage/entry/the_correct_use_of_the_term_nearline2?lang=en (20.8.2015.)

4. TVRDI DISK

Tvrđi disk (engl. *hard disk drive* - *HDD*) je podatkovni centar računala. Gotovo svako računalo i server sadrži jedan ili više tvrdih diskova. Svako glavno računalo (engl. *mainframe*) i superračunalo (engl. *supercomputer*) je najčešće povezano sa stotinama tvrdih diskova. To mnoštvo tvrdih diskova pohranjuje promjenjive digitalne podatke u relativno trajnom obliku.

Diskovna tehnologija je izuzetno brzo napredovala u posljednjih nekoliko godina. Pružajući gotovo neograničenu izravnu pohranu po iznimno niskim cijenama, omogućuje stvaranje visoko kompleksnog softvera bez brige oko ograničenja veličine, a korisnike oslobađa od razmišljanja koje podatke i koliko podataka mogu pohraniti. Iako disk nije središnja jezgra računala, njegov dinamičan rast je odigrao ključnu ulogu u napretku računalnih sustava.

Utjecaj pohrane bazirane na diskovima je, u današnje vrijeme, otišao i izvan računalnih sustava. U obliku ugrađenih uređaja (engl. *embedded devices*) nalazi se u svakodnevnim potrošačkim proizvodima kao što su digitalni video snimači, kamere, glazbeni uređaji, automobilske navigacijske sustave, mobiteli, itd. Osnovna načela diskova su ista, ali takve vrste aplikacija zahtijevaju različite setove zahtjeva i radnih svojstava.

Tvrđi disk je vrlo složen elektro-mehanički sustav koji obuhvaća desetljeća istraživanja širokog spektra raznih disciplina. On je spoj fizike, kemije, znanosti o materijalima, tribologije, elektrotehnike i elektronike, strojarstva, teorije kodiranja, računalne znanosti, te znanosti o proizvodnji.

Tvrđi disk se prvi put pojavio 1956. godine, kao izum američke tvrtke IBM, kapaciteta od samo nekoliko megabajta. Danas se kapaciteti tvrdih diskova mjere u stotinama gigabajta (GB), pa čak i terabajtima (TB). Sve što se nalazi na tvrdom disku je zapravo red bajtova, jer je svaka datoteka (slika, video, tekst, pjesma, ili nešto drugo) zapisana na tvrdi disk kao niz bajtova.¹⁴

Nakon izlaska na tržište, tvrdi disk je postao dominantna tehnologija sekundarne pohrane podataka u računalnim sustavima. Do danas se tvrdi diskovi rabe u svim računalnim sustavima, iako nedavne inovacije u tehnologijama kao što su memorijski diskovi (engl. *solid*

¹⁴ Tvrđi disk. URL: https://bs.wikipedia.org/wiki/Tvrđi_disk(23.8.2015.)

state drives) koji rabe tehnologije *flash* memorija i NAND tehnologije polako potiskuju magnetske diskove iz pojedinih sustava. Unatoč tome, zbog neprekidnog razvoja novih tehnoloških rješenja, tvrdi diskovi su još uvijek najprivlačniji što se tiče odnosa brzine pristupa, kapaciteta i cijene.¹⁵

Primarne karakteristike HDD-a su njegov kapacitet i performanse. Dio kapaciteta tvrdog diska je nedostupan korisniku jer se koristi za datotečni sustav (engl. *file system*) i operativni sustav računala, a ponekad i za ugrađenu zalihost za ispravljanje pogrešaka i oporavak (engl. *recovery*). Performanse su određene vremenom traženja (engl. *seek time*), tj. vremenom do pronalaska traženog cilindra, vremenom promjene glave (engl. *headswitch*), tj. vremenom do pronalaska tražene trake i okretnom latentnošću (engl. *rotational latency*), tj. vremenom dok glava dočeka sektor. Na temelju ovih vrijednosti formulira se vrijeme pristupa podacima (engl. *data access time*) koje je kombinacija vremena traženja, vremena promjene glave i okretne latentnosti.

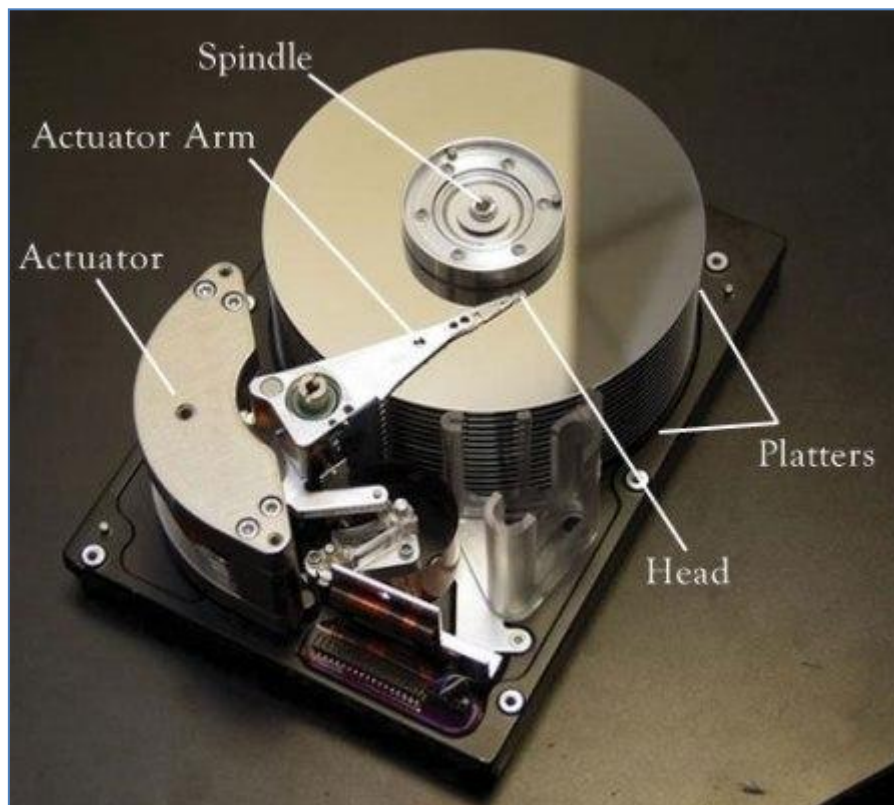
4.1. Građa tvrdog diska

Fizičke osnove magnetskog polja i svojstva feromagnetskih materijala su temelj djelovanja magnetskog diska. Pri upisu podataka, koriste se svojstva tvrdih feromagnetskih tvari da, nakon što su magnetizirane vanjskim magnetskim poljem, ostanu magnetizirane i nakon što vanjsko polje prestane djelovati.

Tvrdi disk koristi kružne plosnate diskove, koji se nazivaju ploče (engl. *platter*). Te ploče su obično od nemagnetskih materijala, najčešće legure aluminijske ili staklene. Premazane su tankim slojem feromagnetskog materijala debljine 10 do 20 nm. Na taj magnetski materijal se još stavlja i premaz ugljika, kao zaštitni sloj. Za magnetski premaz se danas obično biraju legure kobalta, dok su se ranije koristili oksidi željeza, kroma i slično. Ploče imaju rupu u sredini i naslagane su u visinu na osovinu (engl. *spindle*). Ploče se rotiraju velikom brzinom, pogonjene elektromotorom spojenim na osovinu. Brzina okretaja iznosi od 3 do 10 tisuća, a ponekad i više, po minuti. Posebni elektromagnetski uređaji za čitanje i pisanje nazivaju se glave (engl. *heads*) i montirane su na klizačima (engl. *sliders*), te se koriste za upisivanje podataka na disk ili za čitanje podataka sa diska. Klizači su zatim montirani na pokretne ruke.

¹⁵ Tvrdi disk. URL: https://hr.wikipedia.org/wiki/Tvrdi_disk (24.8.2015.)

Sve zajedno je mehanički povezano u jedan sklop i pozicionirano iznad površine ploča mehanizmom za pokretanje (engl. *actuator*). Na današnjim modernim diskovima, udaljenost te glave od površine ploče se mjeri u nanometrima.



Slika 5: Građa tvrdog diska

Preuzeto sa: <http://forums.techarena.in/guides-tutorials/1027662.htm> (24.8.2015.)

Pojedina površina svake ploče na disku, može sadržavati desetke milijardi individualnih bitova podataka. Oni su organizirani u veće blokove zbog praktičnosti, te lakšeg i bržeg pristupa informacijama. Svaka ploča ima dvije glave, jednu na gornjoj strani ploče i jednu na donjoj strani ploči. Stoga tvrdi disk s tri ploče ima šest površina i šest glava. Svaka ploča ima podatke snimljene u koncentričnim kružnicama koje se zovu staze (engl. *tracks*). Sve trake istog rednog broja na svim stranama svih ploča čine jedan cilindar (engl. *cylinder*). Svaka staza se dalje dijeli na manje dijelove koji se zovu sektori (engl. *sectors*), a svaki sektor može sadržavati 512 bajtova podataka.¹⁶

¹⁶ Hard disk drives. URL: <http://www.pcguide.com/ref/hdd/index.htm> (24.8.2015.)

4.2. Pristup podacima na tvrdom disku

Podatke poslane na tvrdi disk i s tvrdog diska tumači disk kontroler (engl. *disk controller*), koji govori tvrdom disku što učiniti i kako pomicati komponente unutar pogona. Kad operativni sustav treba čitati ili zapisivati podatke, pregledava alokacijsku tablicu datoteka (engl. *File Allocation Table – FAT*) tvrdog diska da odredi lokaciju datoteke ili područje dostupno za zapisivanje. Nakon što ju odredi, disk kontroler upućuje mehanizam za pokretanje da pomakne ruke za čitanje i pisanje, te da odabere glavu koja će čitati ili pisati podatke. Budući da su datoteke često raspršene po pločama, glavu treba premjestiti na različite lokacije jedne plohe, te naizmjenice isključivati i uključivati neku drugu glavu zaduženu za drugu plohu, kako bi se pristupilo svim informacijama.

Sva pohrana informacija na tradicionalnom tvrdom disku se vrši magnetski. Nakon završetka navedenih koraka, ako računalo treba čitati podatke s tvrdog diska, čita magnetske polaritete na ploči. Jedna strana magnetskog polariteta je nula, a druga strana je jedan. Čitajući to kao binarne podatke, računalo može razumjeti što su podaci koji se nalaze na ploči.¹⁷

4.3. Lociranje podataka na tvrdom disku

Kada sustav šalje podatke za snimanje, oni se bilježe u grupama od 512 bajtova. Svaka ta grupa se naziva podatkovni blok (engl. *data block*). Bitovi podatkovnih blokova se bilježe uzastopno po stazi. Da bi locirali podatkovni blok na površini, označuju se identifikacijskim brojem. Na razini sustava, svakom podatkovnom bloku je dodijeljena logička adresa bloka (engl. *Logical Block Address – LBA*). Ove adrese blokova, međutim nisu prikladne za pristup podacima na nižoj razini. Za pristup podacima na nižoj razini koristi se broj cilindra, broj glave i broj sektora dodijeljene svakoj logičkoj adresi bloka.

Tvrđi disk može sadržavati više diskova sa podacima snimljenim na obje površine diska. Podacima na pojedinoj površini se pristupa, kako za čitanje tako i za pisanje, korištenjem zasebne glave za tu površinu. Ako tvrdi disk ima n diskova, sadržavati će i $n \times 2$

¹⁷ Hard drive. URL: <http://www.computerhope.com/jargon/h/harddriv.htm>(25.8.2015.)

površina, pa tako i $n \times 2$ glava. Svaka površina diska se identificira po odgovarajućem broju glave. Tvrdi disk s četiri ploče ima osam iskoristivih površina, te glave numerirane od nula do sedam. Postoje deseci tisuća staza na svakoj površini, numeriranih izvana prema unutra, počevši od nule.

Svaka staza je nadalje podijeljena na sektore, koristeći posebne magnetske obrasce napisane na diskovima u trenutku proizvodnje. Ti posebni obrasci su poznati kao slijedni sektori (engl. *servo sectors*), koji dijele staze na jednake segmente, a podaci se upisuju u te segmente. Broj slijednih sektora po stazi je isti na svim površinama u tvrdom disku. U prosjeku jedan tvrdi disk ima 100-200 slijednih sektora.¹⁸

U novije doba se broj sektora razlikuje ovisno o tome jesu li oni u stazi bliže obodu diska ili njegovome centru. Riječ je o zapisivanju bitova u zone (engl. *zone bit recording - ZBR*). Na disku koji se sastoji od koncentričnih staza, fizička duljina staze povećava se što je dalje od sredine. Na diskovima koji koriste ZBR podaci na unutarnjim stazama su zbijeni onoliko gusto koliko im tehnologija određenog tvrdog diska omogućuje, dok su podaci na vanjskim stazama rjeđe zbijeni. Upotrebom ZBR-a disk dijeli staze u više zona. Spremanje više bitova po stazi dovodi do postizanja većeg ukupnog kapaciteta podataka na istom diskovnom području.¹⁹

Također, neki noviji diskovi koriste princip okomitog zapisivanja (engl. *perpendicular recording*). To je metoda digitalnog zapisa na magnetski disk u kojem su bitovi u vertikalnom umjesto u horizontalnom rasporedu sa svrhom zauzimanja manje prostora.²⁰

4.4. Relativna učestalost pogrešaka i rukovanje

Moderni tvrdi diskovi opsežno primjenjuju kôd za ispravljanje pogrešaka (engl. *Error Correcting Code - ECC*), pogotovo Reed-Solomon kôd za ispravljanje pogrešaka. Ove tehnike pohranjuju dodatne bitove, određene matematičkim formulama, za svaki blok podataka. Ti bitovi ne sadrže podatke, već sadrže informacije o podacima koje se koriste za

¹⁸ Al Mamun, A.; Guo, G.; Bi, C.; *Hard Disk Drive: Mechatronics and Control*, Taylor & Francis Group, LLC. Boca Raton, Florida, USA; 2007.

¹⁹ Zone bit recording. URL: https://en.wikipedia.org/wiki/Zone_bit_recording (16.9.2015.)

²⁰ Perpendicular recording. URL: <http://www.pcmag.com/encyclopedia/term/49119/perpendicular-recording> (16.9.2015.)

ispravljanje bilo kakvih problema na koje se može naići pristupajući stvarnim podatkovnim bitovima. Tako ti dodatni bitovi omogućuju da se mnoge pogreške isprave nezapaženo. Dodatni bitovi sami po sebi zauzimaju mjesto na tvrdom disku, ali dopuštaju veću gustoću snimanja bez uzrokovanja pogrešaka koje se ne mogu ispraviti, što rezultira puno većim kapacitetom pohrane.²¹ Što se tiče mjesta koje ti bitovi zauzimaju, uzmimo za primjer tipični tvrdi disk kapaciteta jednog TB-a, koji je podijeljen na sektore veličine 512 bajtova. Takav tvrdi disk ima rezervirani kapacitet od otprilike 93 GB-a za ECC podatke.

U novijim tvrdim diskovima, od 2009. godine, kôd niske gustoće s provjerom pariteta (engl. *low-density parity check code* - LDPC) je zamijenio Reed-Solomonov kôd. LDPC kôdovi omogućuju izvedbu blisku Shannonovom limitu i tako osiguravaju najveću dostupnu gustoću pohrane.²² "Shannonov limit (engl. *Shannon limit*) ili Shannonov kapacitet (engl. *Shannon capacity*) komunikacijskog kanala je teoretska maksimalna brzina prijenosa informacija komunikacijskim kanalom, za određenu razinu buke."²³

Tipični tvrdi diskovi pokušavaju ponovo mapirati podatke iz fizičkog sektora koji se kviri u rezervni fizički sektor, oslanjajući se da će ECC ispraviti pohranjene podatke dok je količina pogrešaka u lošem sektoru dovoljno niska. Značajka tehnologije za samonadziranje, analizu i izvještavanje (engl. *Self-Monitoring, Analysis and Reporting Technology* - S.M.A.R.T.) broji ukupan broj pogrešaka na cijelom tvrdom disku ispravljenih od strane ECC-a i ukupan broj obavljenih ponovnih mapiranja sektora, jer nastanak mnogo takvih grešaka može predvidjeti kvar tvrdog diska.²⁴

Samo mali dio uočenih pogrešaka ostanu kao nepopravljive pogreške. Na primjer, specifikacije poslovnog SAS (engl. *Serial Attached SCSI*) diska (model iz 2014.) procjenjuju jednu neispravljivu pogrešku na svakih 10^{16} bitova.²⁵ Drugi moderni SATA disk iz 2015.

²¹ Error Correcting Code (ECC). URL: <http://pcguide.com/ref/hdd/geom/errorECC-c.html> (28.8.2015.)

²² Galbralth, R.; Oenning, T.; *Iterative Detection Read Channel Technology in Hard Disk Drives*. Whitepaper, Studeni 2008. URL: [http://www.hgst.com/tech/techlib.nsf/techdocs/FB376A33027F5A5F86257509001463AE/\\$file/IDRC_WP_final.pdf](http://www.hgst.com/tech/techlib.nsf/techdocs/FB376A33027F5A5F86257509001463AE/$file/IDRC_WP_final.pdf) (28.8.2015.)

²³ Noisy-channel coding theorem. URL: https://en.wikipedia.org/wiki/Noisy-channel_coding_theorem (28.8.2015.)

²⁴ Self-Monitoring Analysis and Reporting Technology (S.M.A.R.T.) URL: http://www.pcguides.com/ref/hdd/perf/qual/features_SMART.htm (28.8.2015.)

²⁵ Enterprise Performance 15K HDD. URL: <http://www.seagate.com/www-content/product-content/enterprise-performance-savvio-fam/enterprise-performance-15k-hdd/ent-perf-15k-5/en-gb/docs/enterprise-performance-15k-hdd-ds1797-3-1406gb.pdf> (28.8.2015.)

godine u specifikacijama navodi da je stopa pogreške manja od jedne neobnovljive pogreške u čitanju na 10^{15} bitova.²⁶

Najgori tip pogreške je onaj koji ostane nezapažen, a ne mogu ga ni detektirati ni *firmware* diska niti operativni sustav. Te pogreške se nazivaju tiho kvarenje podataka (engl. *silent data corruption*), a neke od njih uzrokuju greške u radu tvrdog diska.

4.5. Redundantno polje nezavisnih diskova

Redundantno polje nezavisnih diskova (engl. *Redundant Array of Independent Disks - RAID*) je tehnika koja koristi više diskova za izgradnju bržeg, većeg i pouzdanijeg diskovnog sustava. Termin je u kasnim 1980-tim uvela skupina istraživača na UC Berkeley (na čelu s profesorima Davidom Pattersonom i Randyjem Katzom, te studentom Garthom Gibsonom).

Izvana, RAID izgleda kao disk: grupe blokova koje se mogu čitati i zapisivati. Iznutra, RAID je složen sustav koji se sastoji od više diskova, memorije (postojane i nepostojane), te jednog ili više procesora koji upravljaju tim sustavom. Hardverski, RAID je vrlo sličan računalnom sustavu, specijaliziran za zadatak upravljanja grupom diskova.

RAID-ovi nude niz prednosti u odnosu na jedan disk. Jedna od prednosti su performanse. Korištenje više diskova paralelno može uvelike ubrzati U/I vrijeme. Druga prednost je kapacitet. Veliki skupovi podataka zahtijevaju velike diskove. Uz to, RAID-ovi poboljšavaju pouzdanost. Ako su podaci pohranjeni na više diskova (bez RAID tehnike), ti podaci su ranjivi na gubitak pojedinog diska. S nekim oblikom redundancije, RAID-ovi mogu tolerirati gubitak diska i nastaviti s radom kao da se nije ništa dogodilo. RAID-ovi pružaju sve te prednosti transparentno sustavima koji ih koriste. Na primjer, računalni sustav vidi RAID kao veliki disk. Ljepota te transparentnosti je ta što omogućuje jednostavnu izmjenu diska unutar RAID konfiguracije bez ikakvih promjena u softveru, a operativni sustav i aplikacije nastavljaju raditi bez izmjena. Zahvaljujući tome, korisnici i administratori mogu koristiti RAID bez brige o kompatibilnosti softvera.

²⁶ WD Re: Datacenter Capacity HDD. URL: <http://www.wdc.com/wdproducts/library/SpecSheet/ENG/2879-800066.pdf> (28.8.2015.)

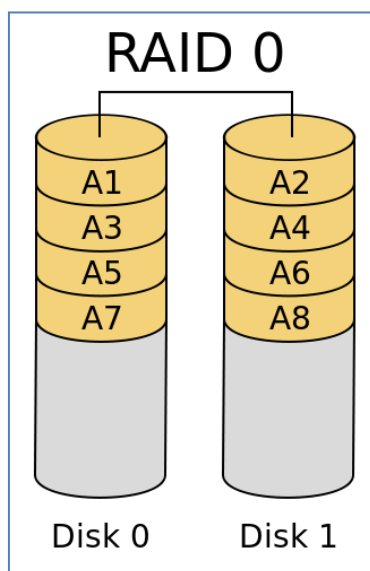
Tri su važna RAID dizajna: RAID razine 0 (segmentacija, cijepanje podataka na više od jednog diska – engl. *striping*), RAID razine 1 (zrcaljenje – engl. *mirroring*), i RAID razine 4/5 (zahilost na bazi pariteta – engl. *parity based redundancy*). Imenovanje svakog od ovih dizajna kao "razina" proizlazi iz pionirskog rada Pattersona, Gibsona i Katza.²⁷

4.5.1. RAID razine 0

RAID razine 0 zapravo nije prava razina RAID-a jer ne koristi zalihost (engl. *redundancy*) za poboljšanje performansi. RAID 0 ima samo segmentaciju, bez zrcaljenja i pariteta.

Kapacitet RAID-a razine 0 je zbroj kapaciteta diskova u setu. Nema dodane zahilosti za rukovanje kvarovima diska. Dakle, kvar jednog diska uzrokuje gubitak cijelog RAID-a razine 0, sa smanjenim mogućnostima za oporavak podataka.

Segmentacija distribuira sadržaj datoteke otprilike podjednako na sve diskove u skupu, pa su operacije čitanja i pisanja na više diskova neizbježno istovremene, a to rezultira poboljšanjem performansi.



Slika 6: RAID razine 0

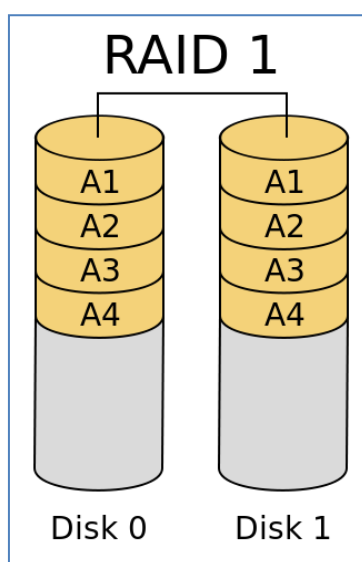
Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID_0.svg (16.9.2015.)

²⁷ Arpaci-Dusseau, R. H.; Arpaci-Dusseau, A. C.; *Operating Systems: Three Easy Pieces*. Arpaci-Dusseau Books, 2015.

Istodobne operacije čine protok većine operacija čitanja i pisanja jednakima protoku jednog diska pomnožen s ukupnim brojem diskova u konfiguraciji. Povećan protok je velika prednost RAID-a razine 0 nad logičkim diskom koji se proteže duž više fizičkih diskova (engl. *spanned volume*).

4.5.2. RAID razine 1

RAID razine 1 se sastoji od zrcaljenja podataka (engl. *data mirroring*), bez pariteta i segmentacije. Podaci se identično zapisuju na dva ili više diskova, čime se stvara zrcalni set diskova. No, ukupan raspoloživi kapacitet je prepolovljen. Tako, na primjer, dva diska od 1 TB svaki, će povezani u RAID 1 dati ukupni kapacitet od 1 TB zbog zapisivanja podataka istovremeno na oba diska.



Slika 7: RAID razine 1

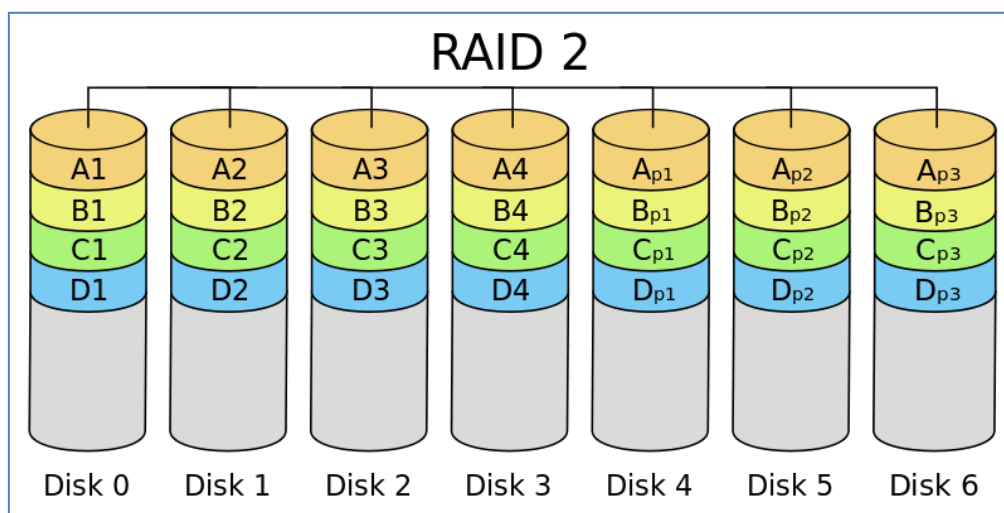
Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID_1.svg (16.9.2015.)

Svaki zahtjev za čitanjem podataka može biti uslužen od bilo kojeg diska u konfiguraciji. Ako se zahtjev emitira svakom disku u setu, uslužuje ga disk koji prvi pristupi podacima, ovisno o njegovom vremenu traženja i okretnoj latenciji, što uvelike poboljšava performanse.

Ako je kontroler ili softver optimiziran, kontinuirani protok iščitavanja se približava zbroju protoka svih diskova u setu, kao i za RAID razine 0. Stvarni protok iščitavanja većine konfiguracija RAID razine 1 je sporiji od najbržeg diska. Protok zapisivanja je uvijek sporiji jer se svaki disk mora ažurirati, a najsporiji disk ograničava performansu zapisivanja. Cijelo polje nastavlja raditi sve dok barem jedan disk funkcionira.²⁸

4.5.3. RAID razine 2

U RAID-u razine 2, tolerancija kvarova se postiže primjenom kôda za ispravljanje pogrešaka (engl. *error-correcting code* – *ECC*) preko svih diskova. Ako se koristi Hammingov kôd, tri redundantna diska se dodaju na svaka četiri podatkovna diska. Korisnički podaci se segmentiraju na razini bita na četiri podatkovna diska. Odgovarajući bitovi s podatkovnih diskova se koriste za izračunavanje tri ECC bita, tako da svaki bit ide na jedan od redundantna tri diska. Čitanje i pisanje se mora raditi paralelno na svim diskovima zbog segmentacije na razini bita.



Slika 8: RAID razine 2

Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID2_arch.svg (16.9.2015.)

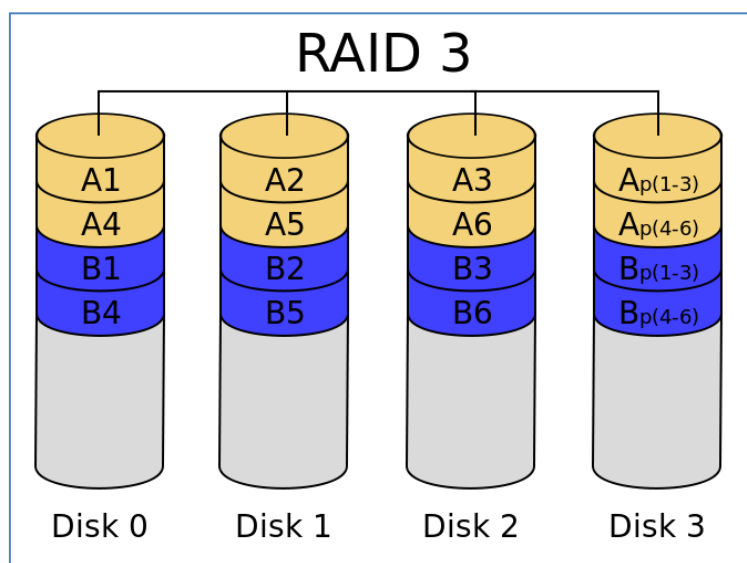
²⁸ Chen, P. M.; Lee E. K.; Gibson, G. A.; Katz, R. H.; Patterson, D. A.; *RAID: High-Performance, Reliable Secondary Storage*. ACM Computing Surveys 26. 1994.

RAID razine 2 je vrlo skupo rješenje za postizanje veće pouzdanosti. Složenije je od zrcaljenja, a nema njegovu fleksibilnost. Nikad nije bio usvojen od industrije za pohranu jer je RAID razine 3 slično, ali jednostavnije i jeftinije rješenje.

4.5.4. RAID razine 3

Ovaj organizacijski koncept je sličan RAID-u razine 2 po tome što se koristi segmentacija na razini bajta, koja je u pravilu identična segmentaciji na razini bita, samo što koristi bajtove kao jedinicu segmentacije.

Umjesto da koristi Hammingov ECC kôd za pružanje korekcija pogrešaka, koristi jednostavnu shemu pariteta kako bi pružio toleranciju kvara jednog od diskova. Dakle, samo jedan redundantan disk –disk pariteta (engl. *parity disk*) treba biti dodan, bez obzira koliko je podatkovnih diskova u konfiguraciji. Jednostavnost RAID-a razine 3, ga čini atraktivnim rješenjem za izradu visoko pouzdanih podsustava pohrane.



Slika 9: RAID razine 3

Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID_3.svg (16.9.2015)

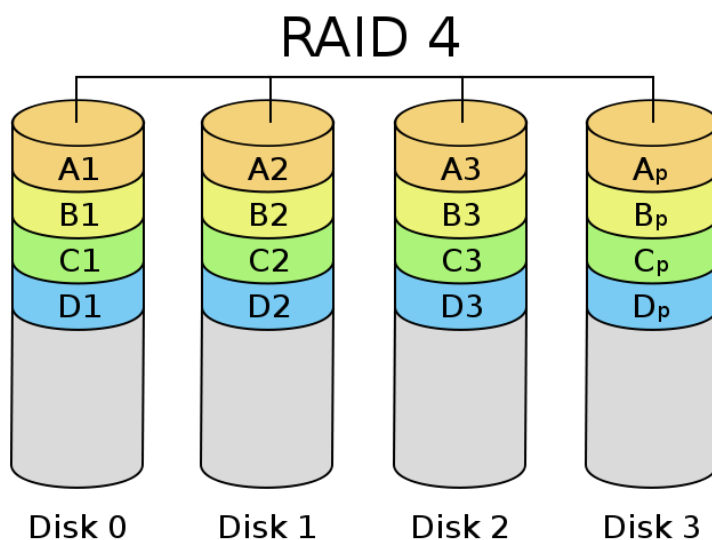
Ova arhitektura je pogodna za aplikacije koje većinom prenose velike količine podataka jer je njen pristup podacima sam po sebi paralelan zbog segmentacije na razini bajta.

Samim time je vrlo popularna kod superračunala. Budući da se diskovima ne može pristupiti individualno kako bi pružili visok broj ulazno-izlaznih operacija po sekundi (engl. *I/O operations per second – IOPS*) za pristupe malim blokovima podataka, RAID razine 3 nije dobro rješenje pohrane za umreženu obradu transakcija (engl. *on-line transaction processing – OLTP*) i druge aplikacije bazirane na bazama podataka.²⁹

4.5.5. RAID razine 4

RAID-ovi razine 4 do 6 koriste nezavisne tehnike pristupa. U polju s nezavisnim pristupom, svaki disk djeluje nezavisno, tako da se različiti U/I zahtjevi mogu zadovoljiti paralelno. Zbog toga, polja s nezavisnim pristupom su pogodna za aplikacije koje imaju velike stope U/I zahtjeva i relativno manje pogodna za aplikacije koje zahtijevaju velike stope prijenosa podataka.

Kao i u drugim RAID sustavima, koristi se segmentacija podataka, ali u RAID-ovima razine 4 do 6, ti segmenti su relativno veliki. U RAID 4 sustavu izračunava se paritetni segment, bit po bit, preko odgovarajućih segmenata na svakom podatkovnom disku i paritetni bitovi se pohranjuju u odgovarajuće segmente na disk pariteta.



Slika 10: RAID razine 4

Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID_4.svg (16.9.2015.)

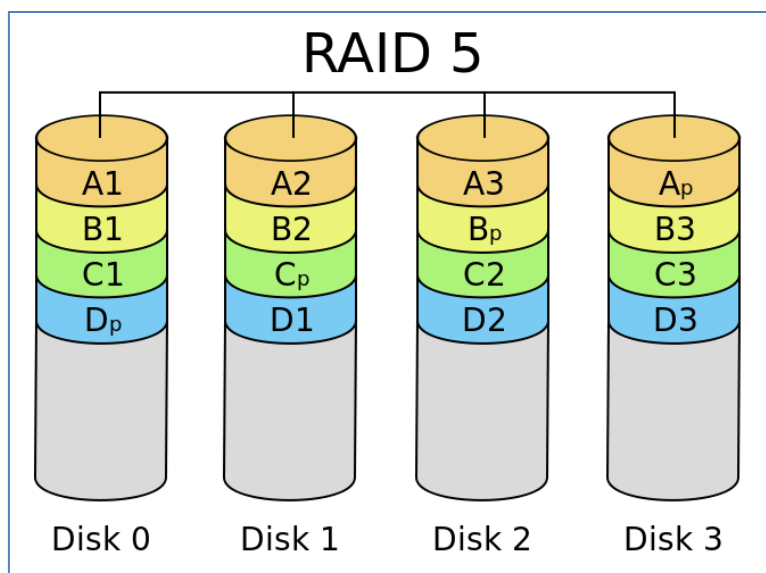
²⁹ n.dj. *Memory systems: Cache, DRAM, Disk*, str. 770-771.

Svaki put kada se podaci zapisuju, softver za upravljanje poljem mora ažurirati, ne samo korisničke podatke, već i odgovarajuće paritetne bitove. Da bi izračunao novi paritet, softver za upravljanje poljem mora čitati stari korisnički segment i stari paritetni segment. Tek tada može ažurirati ta dva segmenta s novim podacima i novo izračunatim paritetom. Dakle svaki segment zapisa uključuje dva čitanja i dva zapisivanja.

U slučaju veće veličine U/I zapisa koje uključuje segmente na svim diskovima, paritet se lagano izračuna koristeći samo bitove novih podataka. Stoga se paritetni disk može ažurirati paralelno s podatkovnim diskovima i ne zahtjeva dodatna čitanja i zapisivanja. U svakom slučaju, svaka operacija zapisivanja uključuje paritetni disk, koji zbog toga može postati usko grlo.

4.5.6. RAID razine 5

RAID 5 je organiziran na sličan način kao i RAID 4. Razlika je u tome što RAID 5 distribuira paritetne segmente po svim diskovima. Tipična raspodjela je shema kružnog dodjeljivanja. Za polje sa n brojem diskova, paritetni segment je na drugom disku za prvih n segmenata, a kasnije se uzorak ponavlja.



Slika 11: RAID razine 5

Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID_5.svg (16.9.2015.)

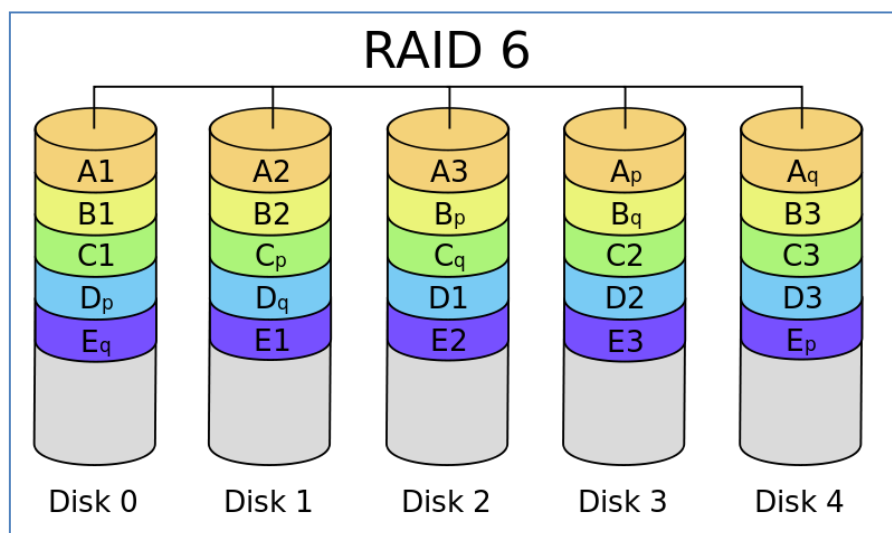
Raspodjela paritetnih segmenata po svim diskovima izbjegava potencijalno U/I usko grlo koje se može naći u RAID-u 4.

4.5.7. RAID razine 6

U RAID 6 konfiguraciji provode se dva različita izračuna pariteta i pohranjuju se u odvojene blokove na različitim diskovima. Dakle, ako korisnički podaci zahtijevaju n diskova, RAID 6 polje se sastoji od $n+2$ diskova.

Dvostruki paritet osigurava toleranciju kvarova i do zatajenja čak dva diska. To čini velike RAID skupine praktičnijima, posebno za sustave visoke dostupnosti, jer visoko kapacitetnim diskovima treba duže vremena za obnovu.

RAID 6 zahtijeva minimalno četiri diska. Kao i kod RAID-a 5, zatajenje jednog diska rezultira smanjenim performansama čitavog polja dok se pokvareni disk ne zamijeni. S RAID 6 poljem, upotrebom diskova s više izvora i više proizvođača, moguće je ublažiti većinu problema povezanih s RAID-om 5. Što su veći kapaciteti diskova i što je veće polje, to je bolje odabrati RAID 6 umjesto RAID 5 konfiguracije.³⁰



Slika 12: RAID razine 6

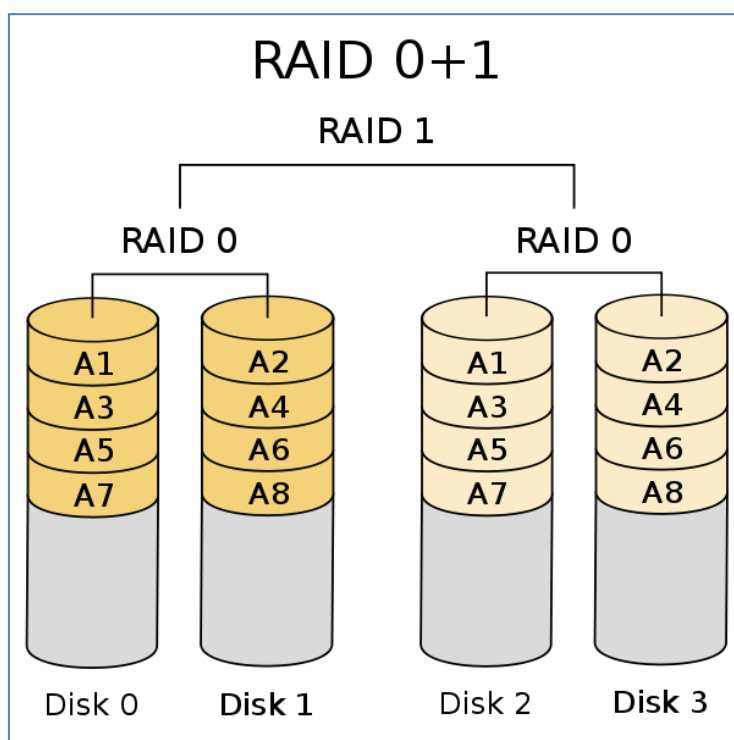
Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID_6.svg (16.9.2015.)

³⁰ n.dj. *Computer Organization and Architecture: Designing for Performance*. str: 194-203.

4.5.8. RAID 0+1

RAID 0+1, koji se još naziva i RAID 01 je RAID razina koja koristi zrcaljenje segmenata, postićući na taj način i replikaciju i razmjenu podataka između diskova. Iskoristivi kapacitet RAID 0+1 polja je isti kao u RAID 1 polju izrađenom od istih tvrdih diskova, u kojem se jedna polovica diskova koristi za zrcaljenje druge polovice. $N/2 \times S_{\min}$, gdje je N ukupan broj diskova i $S_{\min}$ je kapacitet najmanjeg diska u konfiguraciji.

Najmanje četiri diska su potrebna za standardnu RAID 0+1 konfiguraciju, no koriste se i veća polja.³¹



Slika 13: RAID 0+1

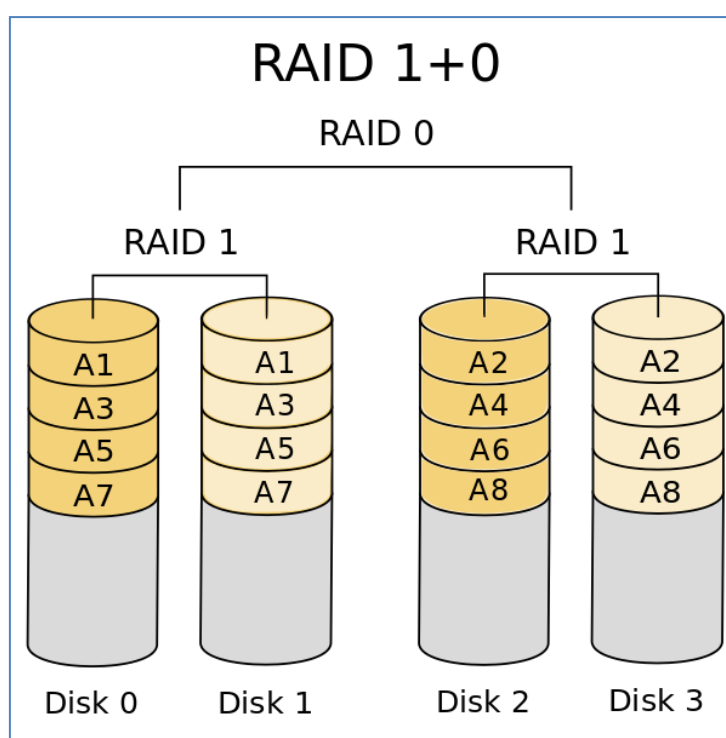
Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID_01.svg (16.9.2015.)

³¹ Nested RAID levels. URL: https://en.wikipedia.org/wiki/Nested_RAID_levels (16.9.2015.)

4.5.9. RAID 1+0

RAID 1+0, koji se još naziva i RAID 10, je slična konfiguracija RAID-u 0+1, s izuzetkom da su dvije korištene standardne RAID razine posložene obrnutim redoslijedom. Dakle, RAID 1+0 je segment zrcaljenja.

RAID 1+0 je RAID 0 zrcaljeno polje, koje može biti dvostruko ili trostruko zrcaljeno i zahtijeva minimalno četiri diska. Konfiguracije sa više od četiri diska su također moguće.



Slika 14: RAID 1+0

Preuzeto sa: https://en.wikipedia.org/wiki/File:RAID_10_01.svg (16.9.2015.)

Prema specifikacijama proizvođača u većini slučajeva RAID 1+0 pruža bolju propusnost i latenciju od svih ostalih razina RAID-a, osim RAID-a 0. Dakle, to je poželjna RAID razina za U/I intenzivne aplikacije kao što su baze podataka, e-mail i web poslužitelji, kao i za bilo koju drugu uporabu koja zahtijeva visoke performanse diska.³²

³² Ibid.

6. PROPADANJE PODATAKA

Opasnosti za integritet podataka uključuju ne samo računalne probleme poput virusa ili hardverske i softverske nekompatibilnosti, kvarova, nedostataka ili propusta, već i prijetnje izvan samog sustava, kao što su nestanci struje, prašina, voda i ekstremna temperatura. Zalutale čestice doslovno mogu promijeniti nulu u jedinicu i obrnuto, uvodeći tako nasumično oštećivanje čak i kada je računalo isključeno. Osim toga fizičko propadanje je stvarna prijetnja. Magnetski diskovi gube zapisanu magnetsku orijentaciju, električki nabijeni mediji postepeno gube naboj, optički mediji se uništavaju jer propada plastika medija ili se oštete. Svaki tvrdi disk, disketa, ili CD, uz dovoljno vremena, postat će pomiješan nered nerazumljivog kôda.

Kvarenje podataka je među najčešćim računalnim pogreškama. Odnosi se na pogreške u računalnim podacima koje uvode neželjene promjene izvornim podacima, a javljaju se tokom čitanja, pisanja, pohrane, prijenosa ili obrade podataka. Kvarenje podataka se događa kada se kôd, namjerno ili slučajno, promijeni iz svog izvornog, ispravnog oblika. Kvarenje može biti sustavno ili nasumično, a čak i mala promjena može iz temelja uništiti program ili učiniti podatak beskorisnim. Računala i sustavi za prijenos ili pohranu koriste niz mjera kako bi osigurali potpun integritet podataka i nedostatak grešaka.³³

Općenito, kada se javi kvarenje podatka, datoteka koja sadrži te podatke će proizvesti neočekivane rezultate kada joj se pokuša pristupiti. Rezultati mogu varirati od manjeg gubitka podataka pa do pada cijelog sustava. Na primjer, ako je iskvarena neka Microsoft Word datoteka, kada se pokuša otvoriti, javit će grešku i neće se moći otvoriti, ili će se uspjeti otvoriti s dijelom iskvarenih podataka. Neki programi, nakon pogreške, mogu dati prijedlog za automatski popravak datoteke, a neki programi ih ne mogu popraviti. Ovisi o razini oštećenja i o ugrađenoj funkcionalnosti aplikacije za upravljanje pogreškama.

³³ Smith, M.; *What is Data Corruption and How to Prevent it*. Srpanj, 2014. URL: <http://www.makeuseof.com/tag/data-corruption-prevent/> (3.9.2015.)



Slika 15: Oštećena .jpeg datoteka u kojoj je većina podataka izgubljena

Preuzeto sa: <http://megasad.com/old/journal/2011-01-01-corrupted-laundry.html> (3.9.2015.)

Tvrdi disk radi tako da diskovni pogon donosi vrlo precizan puls električne energije vrlo malom području diska, mijenjajući tako svoj pohranjeni naboj za označavanje pohranjenog podatka. Ponekad te regije spontano izgube ili promijene svoj naboj, fenomen poznat kao "prevrtanje" (engl. *flipping*). Kada se regija na disku "prevrne" (bitovi promijene svoj naboj), podaci koje ta regija sadrži postaju oštećeni, obrisani ili ostaju nečitljivi. Da bi se naznačila tajanstvena priroda ove degradacije, industrija je razvila termin organskog prizvuka - truljenje podataka (engl. *bit rot*), kako bi objasnili fenomen.³⁴

"Truljenje bitova (engl. *bit rot*) je sporo pogoršanje performansi i integriteta podataka pohranjenih na medijima za pohranu. Također su poznati i nazivi: propadanje bitova (engl. *bit decay*), truljenje podataka (engl. *data rot*), propadanje podataka (engl. *data decay*) ili tiho kvarenje (engl. *silent corruption*)."³⁵

³⁴ Sharwood, S.; *How to beat Bit Rot*. URL: <http://searchstorage.techtarget.com.au/tip/How-to-beat-Bit-Rot> (3.9.2015.)

³⁵ Definition of bit rot. URL: <http://itknowledgeexchange.techtarget.com/discussions/content/bit-rot/> (3.9.2015.)

Postoje dvije vrste kvarenja podataka povezane sa računalnim sustavima. To su detektirano kvarenje podataka i neotkriveno kvarenje podataka.

6.1. Detektirano kvarenje podataka

Otkrivene pogreške mogu biti trajne s gubitkom podataka ili privremene gdje je neki dio sustava u mogućnosti otkriti pogrešku i ispraviti je. O ovom drugom slučaju se ne govori o kvarenju podataka.

Detektirano kvarenje je upravo to, kvarenje koje je otkriveno. To znači da je korisnik svjestan kvarenja. Može se dogoditi da kada se otvori datoteka slike, dio slike bude prekriven artefaktima. Iako je to loše, pruža se prilika da se poduzme akcija. Kada je problem vidljiv, pruža se prilika da se vrati iz sigurnosne kopije (engl. *backup*) i da se detaljnije istraži problem i utvrdi je li kvarenje zbilja bilo nasumično ili je uzrokovano nekim problemom koji se može popraviti.³⁶

Kvarenje podataka se može javiti na bilo kojoj razini u sustavu, od operativnog sustava do medija za pohranu podataka. Moderni sustavi pokušavaju otkriti kvarenje na više slojeva, a zatim obnoviti ili ispraviti kvarenje. To je gotovo uvijek uspješno, ali vrlo rijetko informacija koja dolazi u memoriju sustava bude oštećena i može uzrokovati nepredvidive rezultate.

Kvarenje podataka tijekom prijenosa ima niz uzroka. Prekid prijenosa podataka uzrokuje gubitak podataka. Uvjeti u okolini mogu ometati prijenos podataka, pogotovo kada se radi o bežičnim metodama prijenosa. Npr. teški oblaci mogu blokirati satelitske prijenose. Bežične mreže su podložne smetnjama uređaja kao što su mikrovalne pećnice.

Hardverski i softverski kvarovi su najčešći uzroci gubitka podataka. Pozadinska radijacija, krah glave, i starenje ili trošenje uređaja za pohranu podataka spadaju u hardverske kvarove, dok softverski kvarovi nastaju zbog grešaka u kôdu (engl. *bug*). Kozmičke zrake uzrokuju većinu blagih pogrešaka u DRAM-u.³⁷

³⁶ n. dj. *What is Data Corruption and How to Prevent it*.

³⁷ *Solar Storms: Fast Facts* // Scientific American; srpanj 2008. URL: <http://www.scientificamerican.com/article/solar-storms-fast-facts/> (4.9.2015.)

Najvažnije, detektirano kvarenje je obično manje. Datoteka može biti izgubljena, loš sektor tvrdog diska možda se mora popraviti, ili program više ne radi kako treba. Iako su neugodni, ti slučajevi su obično nadoknadivi i ograničeni opsegom. Čak i kada vode do većih problema, barem daju korisniku vremena za pripremu. Najozbiljniji problemi nastaju kada se kvarenje širi neopaženo.

6.2. Neotkriveno kvarenje podataka

Neotkriveno kvarenje podataka, poznatije pod nazivom tiho kvarenje podataka je najopasnija vrsta pogrešaka jer nema naznaka da su podaci netočni.

Postoje određene vrste pogrešaka u pohrani koje prođu potpuno neprijavljeno i neotkriveno u sustavima za pohranu koje rezultiraju time da aplikacija dohvaća oštećene podatke bez upozorenja, bilježenja, poruke o pogreškama ili obavijesti bilo koje vrste. Taj tip pogrešaka se naziva tiho kvarenje podataka. Ove pogreške je teško otkriti i dijagnosticirati, a dosta su česte u sustavima koji nemaju značajku proširenoga integriteta podataka (engl. *extended data integrity feature*).³⁸

6.2.1. Pogrešno upućen zapis

U nekim slučajevima prilikom zapisa na tvrdi disk, podaci koji su trebali biti zapisani na određenom mjestu, zapravo završavaju zapisani na neko drugo mjesto. Zbog neke greške, disk ne raspoznaje ovu akciju kao pogrešnu, te je izvršava. Kao rezultat toga, pogrešno upućen zapis (engl. *misdirected write*) nije detektiran od strane npr. RAID sustava, jer poduzima akciju samo kada tvrdi disk javi detektiranu pogrešku.

Dakle, nije samo došlo do neotkrivene pogreške, već i do gubitka podataka. Primjerice, blok podataka *C* je trebao biti prepisan preko bloka podataka *A*, ali umjesto toga prepisan je preko bloka *B*. Stoga, blok podataka *B* je izgubljen, a blok podataka *A* još uvijek sadrži krive podatke. Kao rezultat toga podaci su zapisani na pogrešnu lokaciju, jedno

³⁸*Silent data corruption in disk arrays: A solution*. NEC, 2008./2009.; URL: <http://www.necam.com/docs/?id=54157ff5-5de8-4966-a99d-341cf2cb27d3> (5.9.2015.)

područje ima stare, krive podatke, drugo područje ima izgubljene podatke, a ta pogreška nije otkrivena od RAID sustava niti samog tvrdog diska. Pristupi za dohvaćanje podataka bloka *B* ili *C* će rezultirati vraćanjem netočnih podataka bez ikakvog upozorenja.³⁹

6.2.2. *Rastrgan zapis*

U drugim slučajevima, samo neki od sektora koji su trebali biti zapisani zajedno, završe na disku. To se naziva rastrgan zapis (engl. *torn write*) i rezultira blokom podataka koji sadrži djelomično izvorne podatke i djelomično nove podatke. Neki od novih podataka su izgubljeni, a neka čitanja će vraćati stare podatke. Opet, tvrdi disk nije svjestan ove pogreške i vraća uspješan kôd, te npr. RAID opet ne detektira pogrešku. Pristupi za dohvaćanje bloka podataka *B* vraćaju djelomično netočne podatke.⁴⁰

6.2.3. *Oštećenje puta podataka*

Do oštećenja podataka može doći i na njihovom putu do tvrdog diska (engl. *data path corruption*). Put podataka od kontrolera do tvrdog diska se sastoji od nekoliko komponenti. Svaka od tih komponenti se sastoji od manjih komponenti i modula. Iako su podaci zaštićeni detekcijom pogrešaka (kao što je ciklička provjera zalihosti - engl. *Cyclic Redundancy Check* - *CRC*) ili kôdom ispravljanja pokrešaka (*ECC*), dok se kôd prenosi od komponente do komponente, podaci su često izloženi oštećenju dok su još uvijek unutar hardverskih modula i čipova. Na primjer, RAID modul može promijeniti podatke pri izračunavanju pariteta ili neki bitovi mogu biti prevrnuti dok su pohranjeni u predmemoriji kontrolera.

Budući da to mogu biti logičke pogreške, a ne oštećenje uzrokovano kvarom, vjerojatnost da te pogreške budu otkrivene su vrlo male. Uz kombinaciju bugova kôda *firmwarea* i potencijalnog kvara hardvera, tijekom kružnog putovanja od kontrolera do diska i nazad, postoji mogućnost da podaci mogu biti oštećeni bez znanja aplikacije.

³⁹ Krioukov, A.; Bairavasundaram, L. N.; Goodson, G. R.; Srinivasan, K.; Thelen, R.; Arpaci-Dusseau, A.C.; Arpaci-Dusseau, R. H.; *Parity Lost and Parity Regained*. USENIX Conference on File and Storage Technologies. 2008. URL: https://www.usenix.org/legacy/events/fast08/tech/full_papers/krioukov/krioukov.pdf (6.9.2015.)

⁴⁰ Ibid.

6.2.4. Onečišćenje pariteta

Jednom kada dođe do tihog kvarenja, greška može biti složena do točke gdje se izvorni podaci više ne mogu dohvatiti ili uočiti ni od strane RAID sustava.

Kada se podaci zapisuju u sustavu koji koristi RAID razinu s paritetom, novi paritet se izračunava prije nego se podaci zapisuju na disk. RAID sustav očitava podatke i kombinira ih s novim podacima kako bi izračunao novi paritet. U sustavu koji doživljava tiho kvarenje podataka, ova situacija dovodi do nepopravljivog i neotkrivenog gubitka podataka. Kada se iskvareni podaci koriste za izračunavanje novog pariteta, RAID sustav više ne može koristiti paritet da vrati neiskvarene podatke i to nazivamo onečišćenjem pariteta (engl. *parity pollution*).

6.3. Protumjere

Kada se kvarenje podataka ponaša tako da svaki bit podataka ima nezavisno malu vjerojatnost da bude promijenjen, kvarenje podataka se općenito može otkriti pomoću kontrolnih zbrojeva (engl. *checksums*), a često se može i korigirati korištenjem kôdova za ispravljanje pogrešaka. Kontrolni zbroj ili kontrolna suma (engl. *checksum*) je podatak male veličine iz bloka digitalnih podataka u svrhu otkrivanja pogrešaka koje mogu biti uvedene tijekom prijenosa ili pohrane. Kontrolni zbrojevi se koriste za provjeru integriteta podataka, ali se ne bi trebalo oslanjati na njih za provjeru autentičnosti podataka.⁴¹

Ako je otkriveno nepopravljivo oštećenje podataka, mogu se primijeniti postupci kao što su automatski ponovni prijenos (engl. *automatic retransmission*) ili obnova iz sigurnosnih kopija (engl. *backup*). U informacijskim tehnologijama, sigurnosna kopija tj. proces izrade sigurnosnih kopija se odnosi na kopiranje i arhiviranje računalnih podataka, tako da se mogu koristiti za vraćanje izvornih podataka u slučaju gubitka podataka.⁴²

⁴¹ Checksum. URL: <https://en.wikipedia.org/wiki/Checksum> (8.9.2015.)

⁴² Backup. URL: <https://en.wikipedia.org/wiki/Backup> (8.9.2015.)

Određene razine RAID sustava imaju sposobnost za pohranu i procjenu pariteta bitova za podatke preko niza tvrdih diskova i mogu rekonstruirati oštećene podatke uzrokovane kvarom jednog ili više diskova, ovisno o implementiranoj RAID razini.

Mnoge pogreške se mogu otkriti i ispraviti od strane tvrdih diskova upotrebom ECC ili CRC kôdova koji su pohranjeni na disku za svaki sektor. Ako disk otkrije više pogrešaka u čitanju podataka u nekom sektoru, može napraviti kopiju podataka pokvarenog sektora na drugom dijelu diska, tako da ponovo mapira neispravni sektor diska na rezervni sektor bez da u to uključuje operativni sustav (iako to može biti odgođeno do novog zapisivanja u sektor). Ova tiha korekcija se može pratiti pomoću S.M.A.R.T.-a i alata za automatsku provjeru diska za predstojeće kvarove koji su dostupni većini operativnih sustava.

Neki datotečni sustavi (engl. *filesystems*), kao što su Btrfs (*B-tree file system*) i ZFS, koriste kontrolne zbrojeve internih podataka i metapodataka kako bi detektirali tiho kvarenje. Osim toga, ako je oštećenje otkriveno i datotečni sustav koristi interne RAID mehanizme koji pružaju zalihost podataka, takvi datotečni sustavi mogu također rekonstruirati oštećene podatke na transparentan način. Ovaj pristup omogućuje poboljšanu zaštitu integriteta podataka pokrivajući i putove podataka, što je obično poznato kao zaštita podataka s kraja na kraj (engl. *end-to-end data protection*).⁴³

Ispiranje podataka (engl. *data scrubbing*) je još jedan način da se smanji vjerojatnost oštećenja podataka, jer se pogreške diskova uhvate i oporave prije nego se višestruke pogreške akumuliraju i preplave broj paritetnih bitova. Umjesto da se paritet provjerava na svakom čitanju, provjerava se tijekom redovitog skeniranja diska, često kao pozadinski proces niskog prioriteta. Budući da ovaj način aktivira provjeru pariteta, ako korisnik koristi normalan program koji samo čita podatke s diska, paritet neće biti provjeravan osim ako podsustav nema aktiviranu i uključenu provjeru pariteta prilikom čitanja.⁴⁴

⁴³ Zhang, Y.; Rajimwale, A.; Arpaci-Dusseau, A. C.; Arpaci-Dusseau, R.H.; *End-to-end Data Integrity for File Systems: A ZFS Case Study*. USENIX, 2010. URL:

https://www.usenix.org/legacy/events/fast10/tech/full_papers/zhang.pdf (8.9.2015.)

⁴⁴ Data Scrubbing. URL: https://en.wikipedia.org/wiki/Data_scrubbing (8.9.2015.)

7. ZAKLJUČAK

Bilo bi lijepo kada bi se na kraju rada moglo završiti sa optimističnom notom i ponuditi neko tehnološko rješenje koje bi omogućilo aplikacijama da ignoriraju mogućnost kvarova u okruženju dugoročne pohrane. Nažalost, kvarovi su neizbježni. Kako sustavi za pohranu rastu, tako pogreške postaju češće.

Može se primijetiti da je niska razina oštećenja podataka uvijek prisutna i da ima nekoliko izvora. Tehnološki napredak pokušava smanjiti razinu kvarenja podataka, ali je vrlo vjerojatno da neće nikada u potpunosti nestati. Postoji i mogućnost da će novim razvojem hardvera i softvera povremeno doći i do porasta slučajeva i uzroka oštećenja. Potrebno je stalno pažljivo praćenje situacije.

U konačnici, ako se upotrebljavaju odgovarajući mehanizmi za otkrivanje i za uklanjanje oštećenja podataka, izneseni i objašnjeni u ovom radu, integritet podataka se može održati.

Integritet podataka je od velike važnosti u poslovnim aplikacijama, kao što je primjerice bankarstvo, gdje neotkrivena pogreška može ili oštetiti indeks baze podataka ili promijeniti podatke što bi drastično utjecalo na stanje na računu. Važan je i u području arhivistike i dugoročnog očuvanja elektroničkih dokumenata i zapisa kao pretpostavka za potvrđivanje ili utvrđivanje autentičnosti.

Otkrivanje i uklanjanje oštećenja podataka je također važno i kod uporabe šifriranih ili komprimiranih podataka, gdje mala pogreška može učiniti velike skupove podataka neupotrebljivim.

8. LITERATURA

Al Mamun, Abdullah; Guo, GuoXiao; Bi, Chao.; *Hard Disk Drive: Mechatronics and Control*, Taylor & Francis Group, LLC. Boca Raton, Florida, USA; 2007.

Arpaci-Dusseau, Remzi H.; Arpaci-Dusseau, Andrea C.; *Operating Systems: Three Easy Pieces*. Arpaci-Dusseau Books, 2015.

Backup. URL: <https://en.wikipedia.org/wiki/Backup> (8.9.2015.)

Checksum. URL: <https://en.wikipedia.org/wiki/Checksum> (8.9.2015.)

Chen, Peter M.; Lee Edward K.; Gibson, Garth A.; Katz, Randy H.; Patterson, David A.; *RAID: High-Performance, Reliable Secondary Storage*. ACM Computing Surveys 26. 1994.

Computer Data Storage. URL: http://en.wikipedia.org/wiki/Computer_data_storage. (13.8.2015. i 20.8.2015.)

Data Scrubbing. URL: https://en.wikipedia.org/wiki/Data_scrubbing (8.9.2015.)

Definition of bit rot. URL:

<http://itknowledgeexchange.techtarget.com/discussions/content/bit-rot/> (3.9.2015.)

Enterprise Performance 15K HDD. URL: <http://www.seagate.com/www-content/product-content/enterprise-performance-savvio-fam/enterprise-performance-15k-hdd/ent-perf-15k-5/en-gb/docs/enterprise-performance-15k-hdd-ds1797-3-1406gb.pdf> (28.8.2015.)

Error Correcting Code (ECC). URL: <http://pcguide.com/ref/hdd/geom/errorECC-c.html> (28.8.2015.)

Federal Standard 1037C – Telecommunications: Glossary of Telecommunication Terms. URL: <http://www.its.bldrdoc.gov/fs-1037/fs-1037c.htm> (21.8.2015.)

Galbralth, Richard; Oenning, Travis; *Iterative Detection Read Channel Technology in Hard Disk Drives*. Whitepaper, Studeni 2008. URL: [http://www.hgst.com/tech/techlib.nsf/techdocs/FB376A33027F5A5F86257509001463AE/\\$file/IDRC_WP_final.pdf](http://www.hgst.com/tech/techlib.nsf/techdocs/FB376A33027F5A5F86257509001463AE/$file/IDRC_WP_final.pdf) (28.8.2015.)

Hard disk drive. URL: https://en.wikipedia.org/wiki/Hard_disk_drive (24.8.2015.)

Hard disk drives. URL: <http://www.pcguide.com/ref/hdd/index.htm> (24.8.2015.)

Hard drive. URL: <http://www.computerhope.com/jargon/h/harddriv.htm> (25.8.2015.)

Jacob, Bruce; Ng, Spencer W.; Wang, David T.; *Memory Systems: Cache, DRAM, Disk*. Elsevier Inc., Burlington, Massachusetts, USA, 2008.

Krioukov, A.; Bairavasundaram, L. N.; Goodson, G. R.; Srinivasan, K.; Thelen, R.; Arpaci-Dusseau, A.C.; Arpaci-Dusseau, R. H.; *Parity Lost and Parity Regained*. USENIX Conference on File and Storage Technologies. 2008. URL: https://www.usenix.org/legacy/events/fast08/tech/full_papers/krioukov/krioukov.pdf (6.9.2015.)

Memory. URL: <http://www.yourdictionary.com/memory>. (14.8.2015.)

Microsoft Computer Dictionary: Fifth Edition. Microsoft Press, Redmond Washington, SAD, 2002.

Nested RAID levels. URL: https://en.wikipedia.org/wiki/Nested_RAID_levels (16.9.2015.)

Noisy-channel coding theorem. URL: https://en.wikipedia.org/wiki/Noisy-channel_coding_theorem (28-8-2015.)

Patterson, David A.; Hennessey, John L. *Computer Organization and Design: The Hardware/Software Interface*. Fifth Edition. Elsevier Inc., Waltham, Massachusetts, USA, 2014.; str. 372-379.

Pearson, Tony; *The Correct Use of the term Nearline*. URL: https://www.ibm.com/developerworks/community/blogs/InsideSystemStorage/entry/the_correct_use_of_the_term_nearline2?lang=en (20.8.2015.)

Perpendicular recording. URL: <http://www.pcmag.com/encyclopedia/term/49119/perpendicular-recording> (16.9.2015.)

Self-Monitoring Analysis and Reporting Technology (S.M.A.R.T.) URL: http://www.pcguide.com/ref/hdd/perf/qual/features_SMART.htm (28.8.2015.)

Sharwood, S.; *How to beat Bit Rot*. URL: <http://searchstorage.techtarget.com.au/tip/How-to-beat-Bit-Rot> (3.9.2015.)

Silent data corruption in disk arrays: A solution. NEC, 2008./2009.; URL: <http://www.necam.com/docs/?id=54157ff5-5de8-4966-a99d-341cf2cb27d3> (5.9.2015.)

Smith, Matt; *What is Data Corruption and How to Prevent it.* Srpanj, 2014. URL: <http://www.makeuseof.com/tag/data-corruption-prevent/> (3.9.2015.)

Solar Storms: Fast Facts // Scientific American; srpanj 2008. URL: <http://www.scientificamerican.com/article/solar-storms-fast-facts/> (4.9.2015.)

Stallings, William. *Computer Organization and Architecture: Designing for Performance.* Eight Edition. Pearson Prentice Hall, Upper Saddle River, New Jersey, USA, 2010.

Stančić, Hrvoje. Arhivsko gradivo u elektroničkom obliku: mogućnosti zaštite i očuvanja na dulji vremenski rok. *Arhivski vjesnik*, No. 49, prosinac 2006.; URL: http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=9508. (12.8.2015.)

Storage hierarchy. URL: http://www.ts.avnet.com/uk/products_and_solutions/storage/hierarchy.html. (13.8.2015.)

Storage. URL: <http://www.yourdictionary.com/storage>.(14.8.2015.)

Thibodeau, Kenneth. Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. *The State of Digital Preservation: An International Perspective*, Council on Library and Information Resources (CLIR), Washington, D.C., SAD, srpanj 2002.; str. 4-15.

Tvrđi disk. URL: https://bs.wikipedia.org/wiki/Tvrđi_disk (23.8.2015.)

Tvrđi disk. URL: https://hr.wikipedia.org/wiki/Tvrđi_disk (24.8.2015.)

WD Re: Datacenter Capacity HDD. URL: <http://www.wdc.com/wdproducts/library/SpecSheet/ENG/2879-800066.pdf> (28.8.2015.)

Zhang, Y.; Rajimwale, A.; Arpaci-Dusseau, A. C.; Arpaci-Dusseau, R.H.; *End-to-end Data Integrity for File Systems: A ZFS Case Study.* USENIX, 2010. URL: https://www.usenix.org/legacy/events/fast10/tech/full_papers/zhang.pdf (8.9.2015.)

Zone bit recording. URL: https://en.wikipedia.org/wiki/Zone_bit_recording (16.9.2015.)

SAŽETAK

U dobu kada su digitalni podaci dominantni, a pohrana velikog kapaciteta je pristupačna, svaki digitalni podatak se isplati sačuvati. Ovaj rad pruža uvid koliko se ti podaci mogu dugoročno očuvati i koje se sve prijetnje javljaju na samoj fizičkoj razini zapisa. Kroz usporedbu memorije i pohrane, te organizacije pohrane i njene hijerarhije, nudi rješenje kako predstaviti korisniku što je više moguće pohrane u najjeftinijoj tehnologiji, pružajući pristup brzinom koju nudi najbrža memorija. Opisana je građa i način rada tvrdog diska. Opisani su i razni sustavi spajanja više tvrdih diskova u redundantna polja nezavisnih diskova, poznatijim kao RAID konfiguracije, od kojih svaka nudi određene prednosti i nedostatke što se tiče performansa i zaštite podataka. Posljednje poglavlje nudi pregled raznih grešaka i oštećenja koja se mogu dogoditi podacima na samoj razini zapisa na medij, te nekoliko protumjera.

KLJUČNE RIJEČI: hijerarhija memorije, radna memorija, pohrana, tvrdi disk, redundantno polje nezavisnih diskova, RAID, propadanje podataka, kvarenje podataka

Long-term data preservation on a physical level of inscription

SUMMARY

In an age when digital data is dominant and storage of a large capacity is affordable, digital data is worth preserving. This thesis provides insight into how the data can be preserved in the long-term and which threats are occurring on the physical level of records. Through the comparison of memory and storage, and storage organization and its hierarchy, it offers a solution to present the user as much as possible storage capacity in the cheapest technology, providing access at the speed of the fastest memory. The structure and operation of the hard drive is described. Various systems for connecting multiple hard drives in a redundant array of independent disks, known as RAID configurations, are described each of which offers certain advantages and disadvantages when it comes to performance and data protection. Last chapter provides an overview of various errors and types of corruption that can happen to the data on the level of inscription on the media. Several countermeasures are provided as well.

KEY WORDS: memory hierarchy, RAM, storage, hard disk, redundant array of independent disks, RAID, bit rot, data decay, data degradation, data corruption