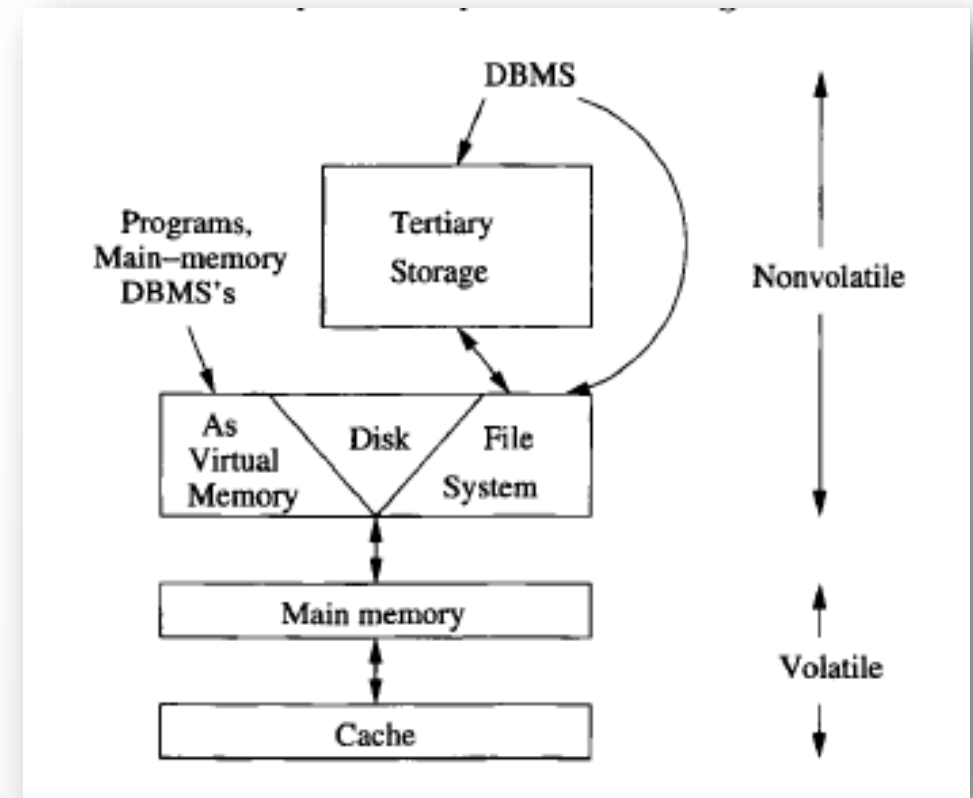


Skladišta podataka

STORAGES

Hijerarhija memorija

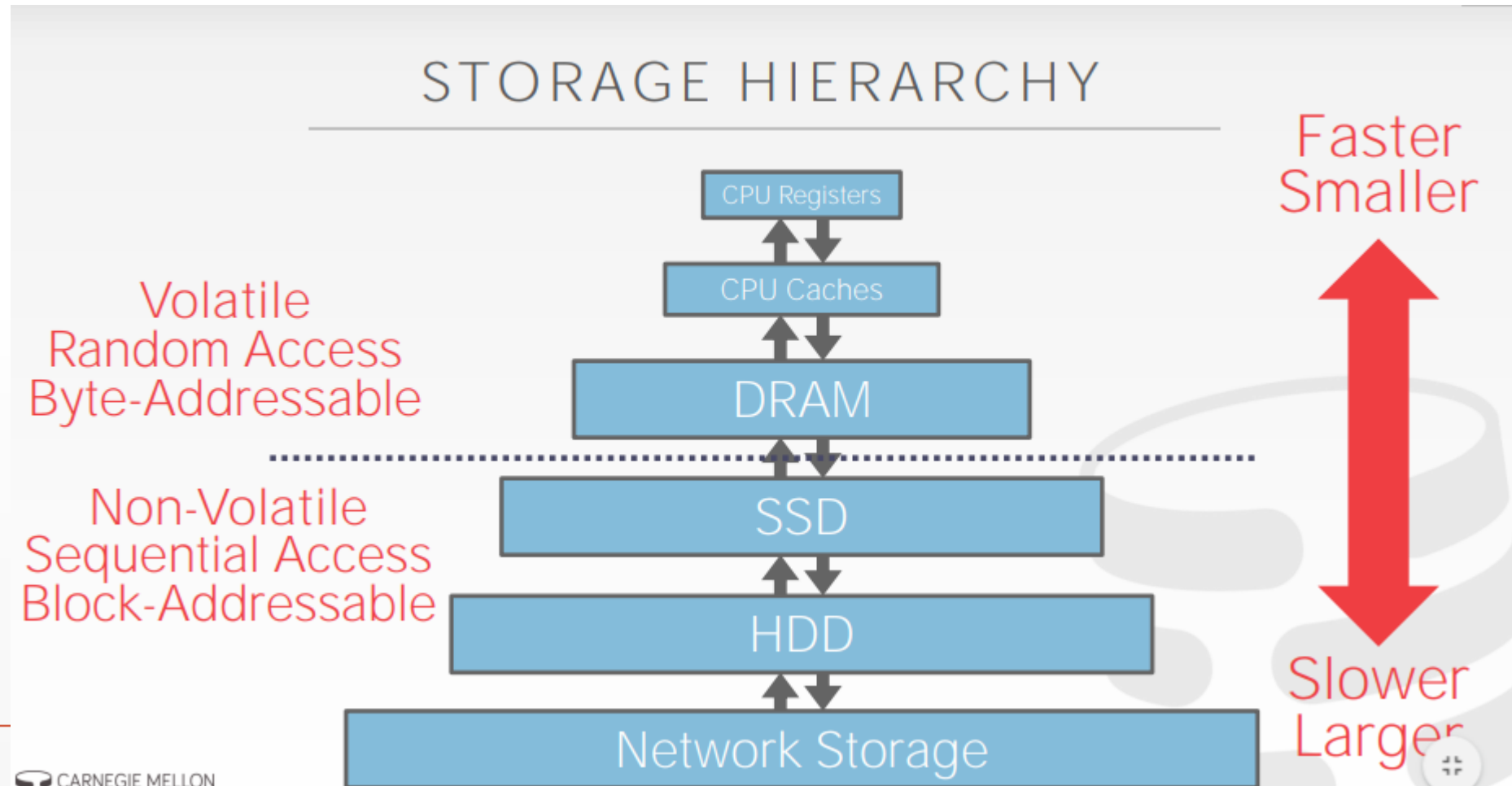
- Razvrstavanje memorijskih komponenti
 - kapacitet – raspon namanje 7 redova veličine
 - brzina pristupa – raspon namanje 7 redova veličine
 - cena po bajtu – 3 reda veličine
 - trajnost:
 - Privremena (Volatile)
 - Trajna (Nonvolatile)



Brzina pristupa

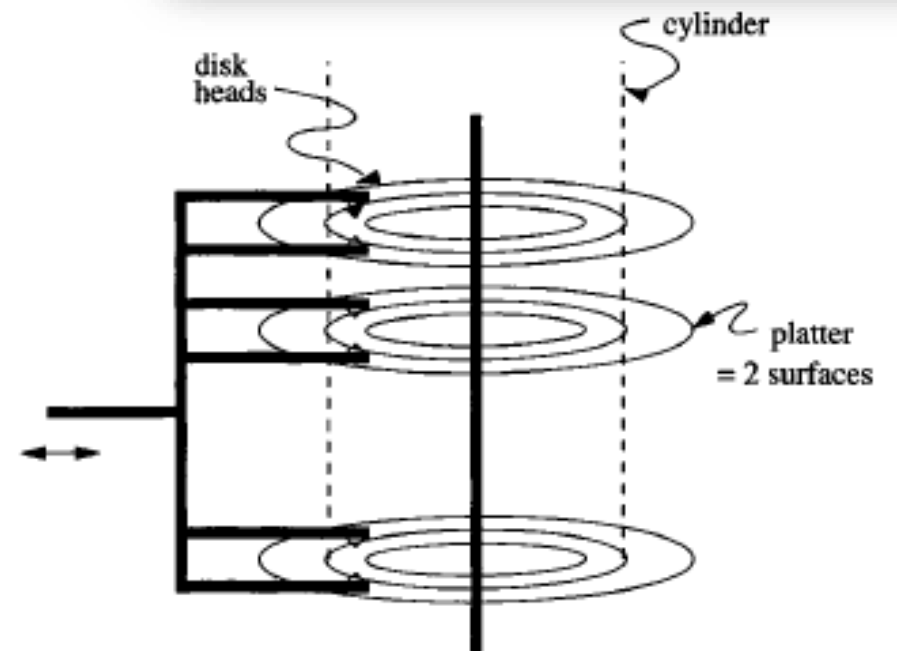
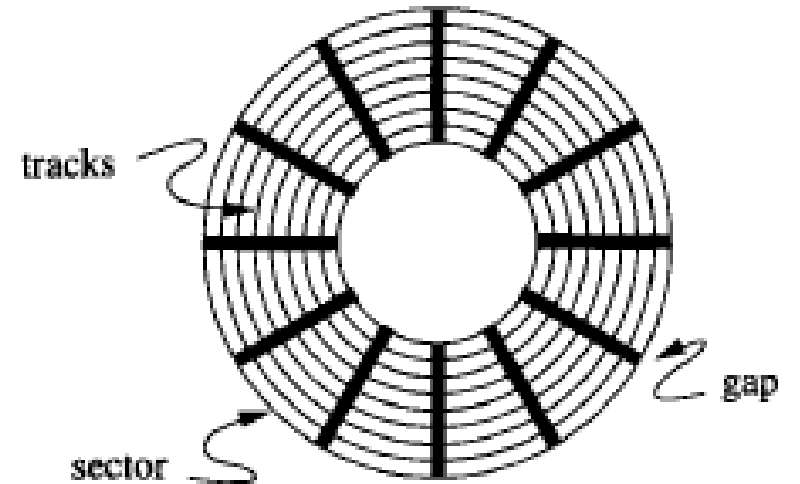
ACCESS TIMES		
0.5 ns	L1 Cache Ref	← 0.5 sec
7 ns	L2 Cache Ref	← 7 sec
100 ns	DRAM	← 100 sec
150,000 ns	SSD	← 1.7 days
10,000,000 ns	HDD	← 16.5 weeks
~30,000,000 ns	Network Storage	← 11.4 months
1,000,000,000 ns	Tape Archives	← 31.7 years

Hijerarhija skladišta



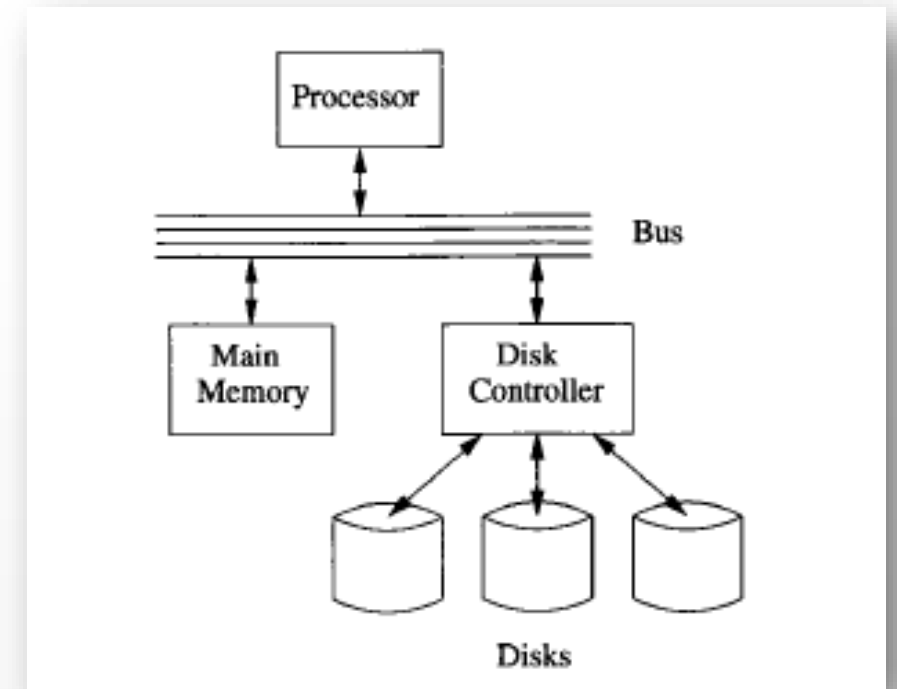
Hard disk

- **Staze** (traks) - 50K-100K po disku
- Staza podeljena na **sektore** izmedju kojih su nenamagnetisani delovi (**gaps – 10%**).
 - Sektor je najmanja jedinica podataka koja može biti upisana ili pročitana (sa stanovišta pisanja i čitanja sa diska, sektori su nedeljivi)
 - Tipična veličina 512 bajtova
 - 500 -1000 na unutrašnjim stazama i 1000-2000 na spoljašnjim



Komunikacija sa ostatkom sistema

- Disk kontroler – procesor sposoban da kontroliše rad jednog ili više disk uređaja.
 - Inicira pozicioniranje glava na odgovarajući cilindar
 - Odlučuje o tome kada je glava pozicionirana na početak sektora koji treba da bude pročit
 - Inicira prenos podataka iz/u sektor
 - Reguliše potencijalno baferisanje čitave/ih staza u lokalnu memoriju kontrolera
 - Odgovoran za pridruživanje checksum-a svakom sektoru, tj. postavljanje mehanizama za proveru ispravnosti prenetih podataka



Familije disk interfejsa

- ATA (AT adaptor) range of standards
 - SATA (Serial ATA)
 - SCSI (Small Computer System Interconnect) range of standards
 - SAS (Serial Attached SCSI)
-

Komunikacija van kontrolera

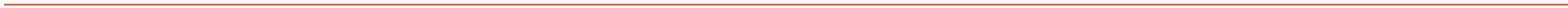
- Postoje i diskovi koji su vezani direktno za računarski sistem.
 - Storage Area Networks (SAN) – veliko broj diskova povezanih sa određenim brojem servera preko mreže visokih brzina.
 - Network Attached Storage (NAS) je umreženo skladište koje poseduje fajl sistem realizovan upotrebom protokola mrežnih fajl sistema.
-

Karakteristike diskova

- **Access time / Vreme pristupa** – vreme koje se protekne od trenutka kada je čitanje/pisanje zahtevano do trenutka kada transfer započne. Sastoji se od:
 - **Vreme pretrage/pronalaženja** – vreme potrebno za pozicioniranje glave čitača
Prosečno vreme pozicioniranja je $\frac{1}{2}$ od najgoreg vremena pozicioniranja.
4-10 ms
 - **Rotaciono kašnjenje** – vreme potrebno da se početak traženog sektora nađe ispod glave
 $\frac{1}{2}$ od najgoreg vremena
4-11 ms (5400 - 1500)
 - **Data-transfer rate / Stopa prenosa podataka** – stopa prenosa na ili sa diska
 - 25-100 MB/s
 - Više diskova može deliti isti kontroler, pa je i njegova stopa prenosa važna
SATA: 150 MB/s, SATA-II 300MB/s
Ultra 320 SCSI 320 MB/s, SAS 3-6 GB/s
-

Karakteristike diskova

- **Srednje vreme otkaza (MTTF)** – prosečno vreme u kojem se očekuje da diska radi bez bilo kakvih otkaza



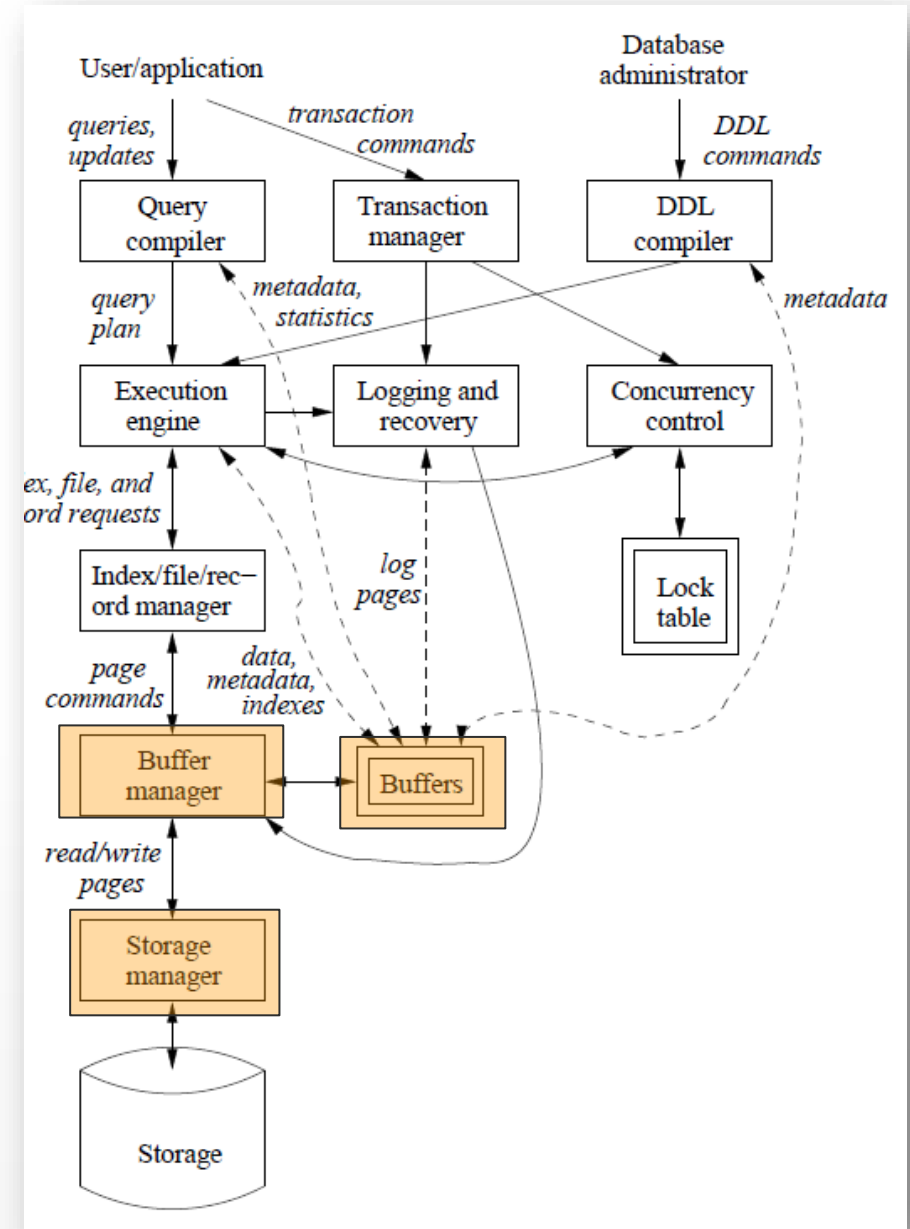
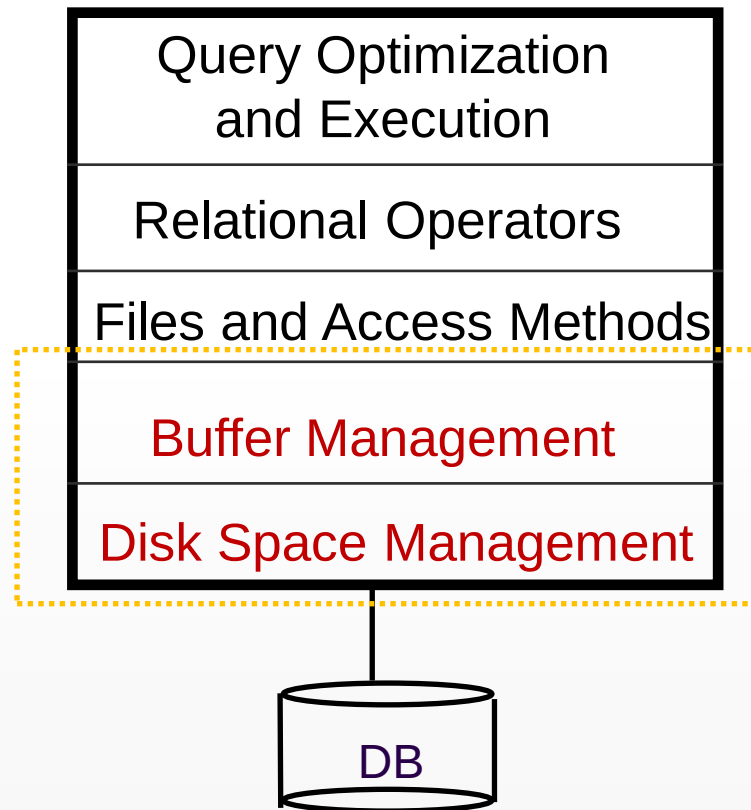
Ciljevi dizajna DBMS-a

- Omogućiti DBMS-u da upravlja bazama podataka koje prevazilaze kapacitete dostupne memorije.
 - Čitanje/pisanje na disk su skupe, pa je od presudnog značaja optimizacija broja takvih operacija.
 - Direktan pristup podacima na disku je značajno sporiji od sekvencijalnog čitanja.
 - Tradicionalni DBMS-ovi su dizajnirani da maksimizuju zastupljenost sekvencijalnog čitanja.
 - Algoritmi za optimizaciju teže da smanje broj direktnih pristupa stranicama tako što se trude da speštaju podatke u susedne blokove.
-

Upravljanje prostorom na disku

Disk space management / Storage manager komponenta

DBMS kontekst



OS fajl sistem v.s. DBMS menadžer prostora

- Uloga menadžera prostora jeste upravljanje prostorom na disku, pri čemu se vodi evidencija o zauzetosti svih blokova koji su na raspolaganju.
 - Upravljanje skladištem može biti izvedeno na dva načina:
 - Oslanjanjem na OS fajl sistem
 - Razvojem upravljača diskom od nule
 - Dvoslojno – menadžer prostora koji sarađuje sa OSom
 - O čemu treba voditi računa:
 - DBMS treba da bude što nezavisniji od platforme da bi bio prenosiv
 - Ograničenja u veličini jednog fajla od strane OS fajl sistema
 - Tipični OS fajlovi se ne mogu protezati na više disk uređaja, što je često potrebno kada su DBMSovi u pitanju
-

File storage

- DBMS čuva beleži podatka u jednom ili više fajlova (OS). Jedan fajl na fajl sistemu se sastoji iz više blokova, ali ih kao takve vidi u DMBS-u vidi samo Menadžer skladišta.
 - Svaki DBMS ima sopstvenu filizofiju organizacije baze po fajlovima.
 - OS ne zna ništa o sadržaju tih fajlova.
-

Upravljanje prostorom na disku

- **Storage Manager (menadžer skladišta)** – najniži nivo u DBMS arhitekturi
- Podržava **koncept strane (page)** kao jedinice podataka i obezbeđuje komande za alociranje i oslobađanje, kao i čitanje i pisanje u stranu.

the size of the page = size of a disk block

- Svaka strana ima jedinstven identifikator. Menadžer skladišta mapira identifikatore strana na fizičke lokacije na disku (blokove).
- Podaci se sa diska dopremaju u radnu memoriju u blokovima

Veličina bloka – od 512 bajtova do nekoliko kilobajta

- Manji blokovi – više transfera
 - Veći blokovi – više protraćenog prostora
 - Tipična veličina - 4 do 64 kilobajta
-

Veličina stranica

DATABASE PAGES

There are three different notions of "pages" in a DBMS:

- Hardware Page (usually 4KB)
- OS Page (usually 4KB)
- Database Page (1-16KB)

By hardware page, we mean at what level the device can guarantee a "failsafe write".

1KB



4KB



ORACLE®

8KB



16KB

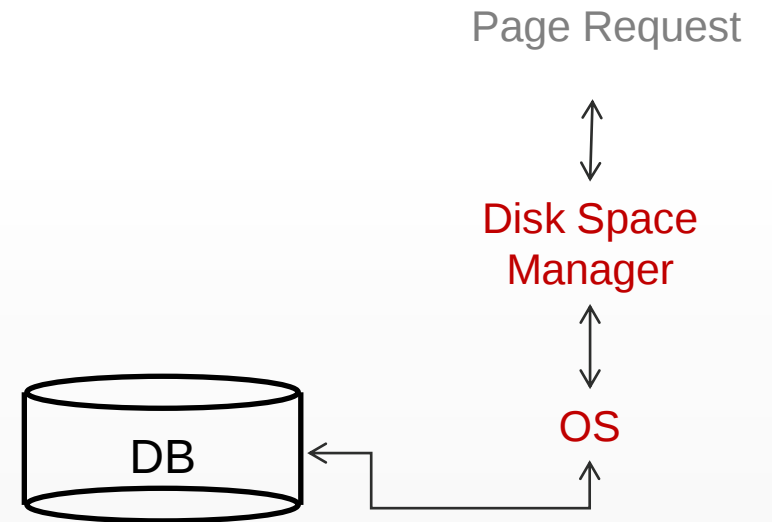


Upravljanje prostorom na disku

- Storage manager sakriva detalje hardvera i operativnog sistema koji su u sloju ispod i omogućava višim slojevima komponenti DBMSa da posmatra podatke u kontekstu kolekcije strana.
 - Zadužen za održavanje fajlova baze podataka.
 - Vodi beleške o:
 - Tome koji su blokovi u upotrebi
 - održavajući listu slobodnih blokova ili
 - održavajući bitmapu sa po jednim bitom za svaki blok
 - Koje strane su na kojim blokovima.
-

Prenos podataka između memorija

- Viši nivoi DBMSa se obraćaju menadžeru skladišta da:
 - alocira/oslobodi stranu
 - čita/piše u stranu
- Najpovoljnija situacija – zahtevane stranice su smeštene redom na disku
- O smeštanju stranica i slobodnom prostoru vodi računa menadžer skladišta, dok o tome viši nivoi DBMSa ne znaju ništa.



Pristup bloku na disku

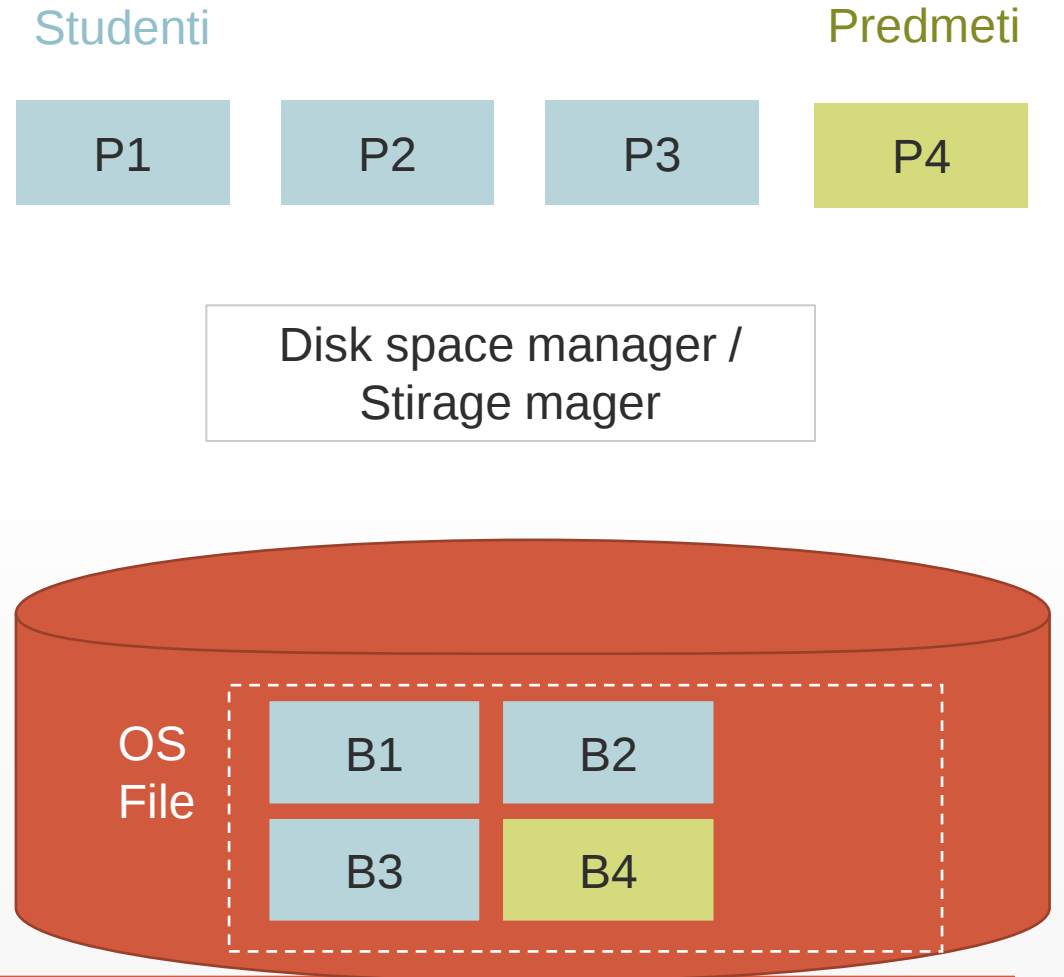
- Vreme pristpa: vreme pozicioniranja glave + rotaciono kašnjenje + vreme transfera
0.3-10 ms 0 – 6 ms 0.008 ms po 8K bloku
 - Smanjenje I/O cene – smanjiti vreme potrošeno na pozicioniranje i rotaciju
 - Blokovi u fajlu bi trebali da budu smešteni sekvencijalno, u uzastopne/susedne blokove
Koncept susednog bloka
 - Blokovi se smeštaju na istu stazu, pa
 - Na isti cilindar, pa
 - Na susedne cilindre
-

DBMS 'virtuelni' fajlovi

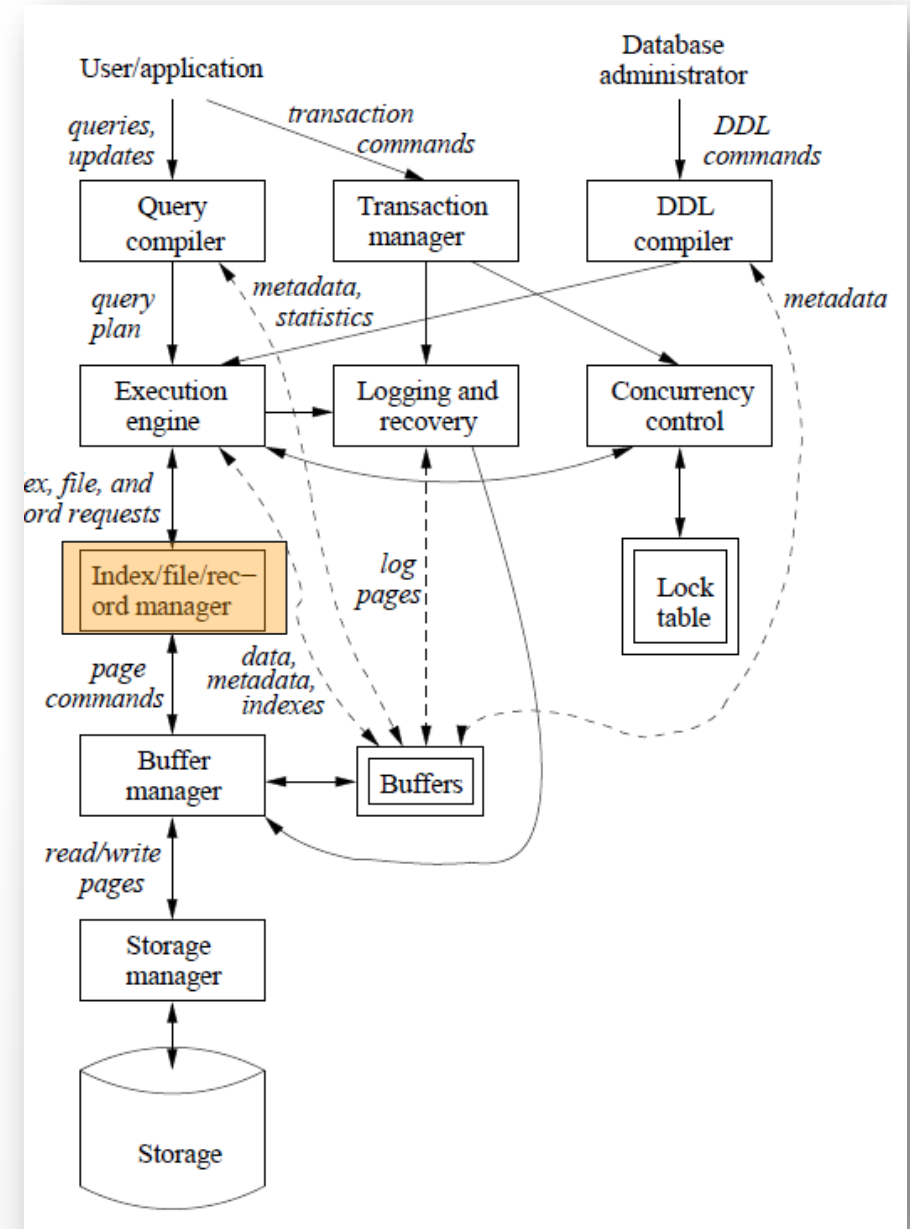
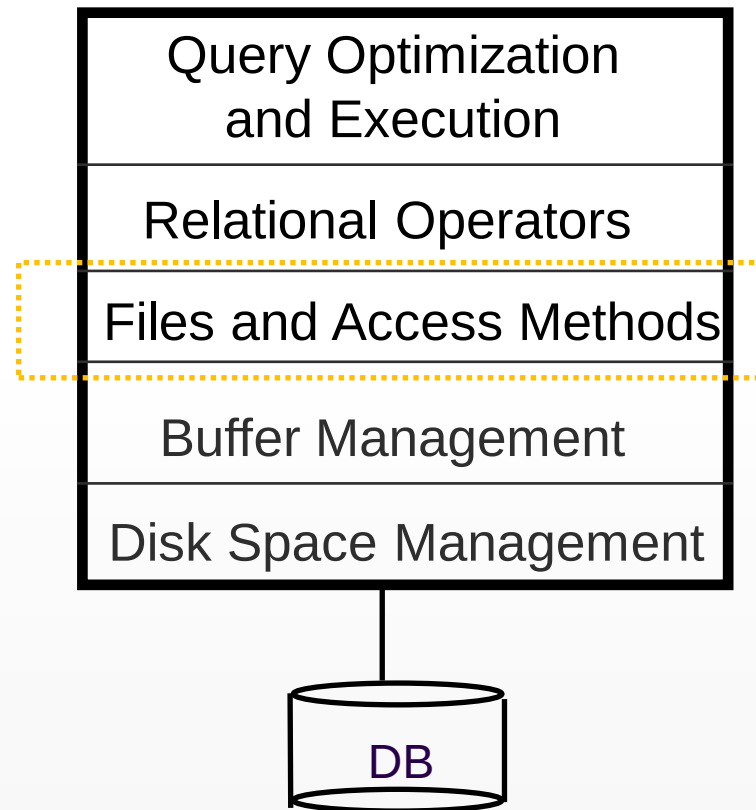
File of records

File of records

- Sve komponente koje rade sa podacima ih posmatraju u kontekstu kolekcija slogova, tj. **slogovnih fajlova** (*file of records*).
- Slogovni fajl obično sadrži slogove jedne tabele.
- Slog – n-torka polja (*fields*).
- Slogovni fajlovi su logičke organizacione jedinice i predstavljaju kolekciju strana.
- **Baza podataka** čini kolekciju slogovnih fajlova.



DBMS kontekst



Održavanje slogovnih fajlova

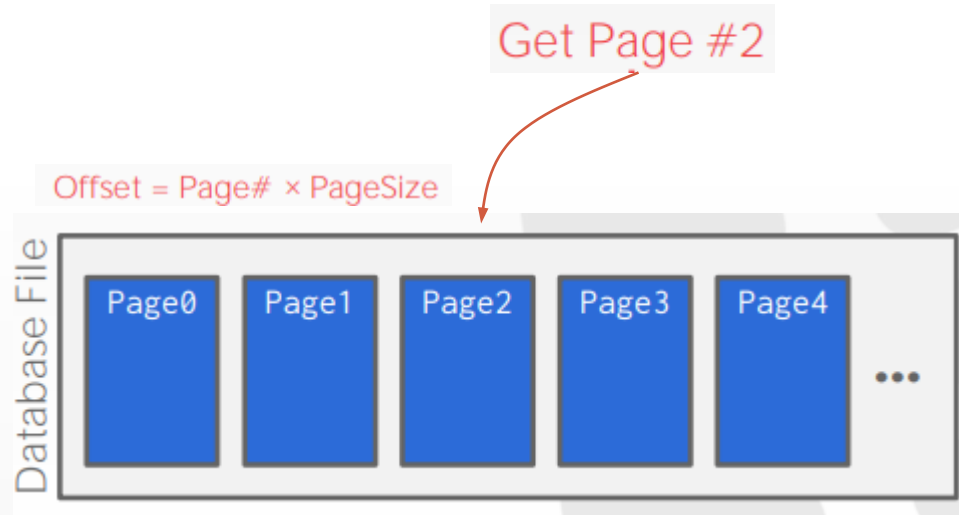
- Menadžer fajlova (index/file/record manager) formira i održava fajl kao kolekciju strana, u kojima se nalaze slogovi.
- Svaka strana
 - je kolekcija slotova, od kojih svaki sadrži po jedan slog
 - svaka strana u fajlu je iste veličine
- Svaki slog (n-torka, record) se identifikuje sledećim parom
(pageID, slot number)

Koji se naziva identifikatorom sloga, record id (RID)

Neuređeni (heap) fajlovi

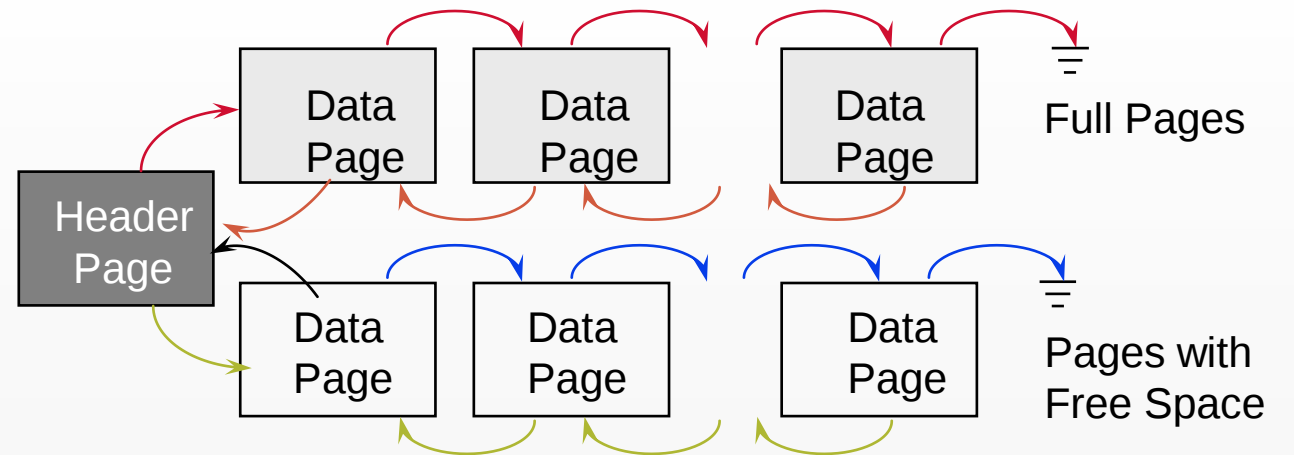
- Najjednostavnija DB fajl struktura ne podrazumeva nikakvo uređivanje slogova unutar fajla. (neka DB fajl bude naziv za File of records)
 - Fajlovi se šire i smanjuju, pa se prema optremi alociraju i dealociraju strane.
 - Da bi se obezbedile operacije sa slogovima, potrebno je voditi evidenciju o
 - Stranama u fajlu
 - Slobodnom prostoru na strani
 - Slogovima unutar strane
 - Održavanje informacija o slobodnim slotovima na stranama.
 - Treba voditi evidenciju o:
 - Slobodnom prostoru unutar strane i
 - Stranama koje imaju slobodnog prostora
-

Kako voditi evidenciju stranama koje pripadaju DB fajlu



Heap fajl implementiran kao lista

- Da bi obezbedili održavanje informacija o slobodnom prostoru, heap fajl je moguće implementirati kao **dve dostruko povezane liste**, gde je jedna **lista popunjenih strana**, a druga **lista strana sa slobodnim prostorom**.
- Za sve fajlove DBMS pamti informaciju o prvoj strani fajla, tzv. **header strani**. Tako održava tabelu koja sadrži parove
(heap_file_name, page_Laddr)



Heap fajl implementiran kao direktorijum strana

- Direktorijum je kolekcija strana.
- Svaki zapis (entry) u direktorijumu identifikuje stranu ili kolekciju strana heap fajla.
- Evidencija o slobodnom prostoru se izvodi tako što se svakom zapisu u direktorijumu pridružuje:
 - bit koji ukazuje na to da li strana ima ili ne slobodnog prostora
 - broj kojim je predstavljena količina slobodnog prostora na strani

