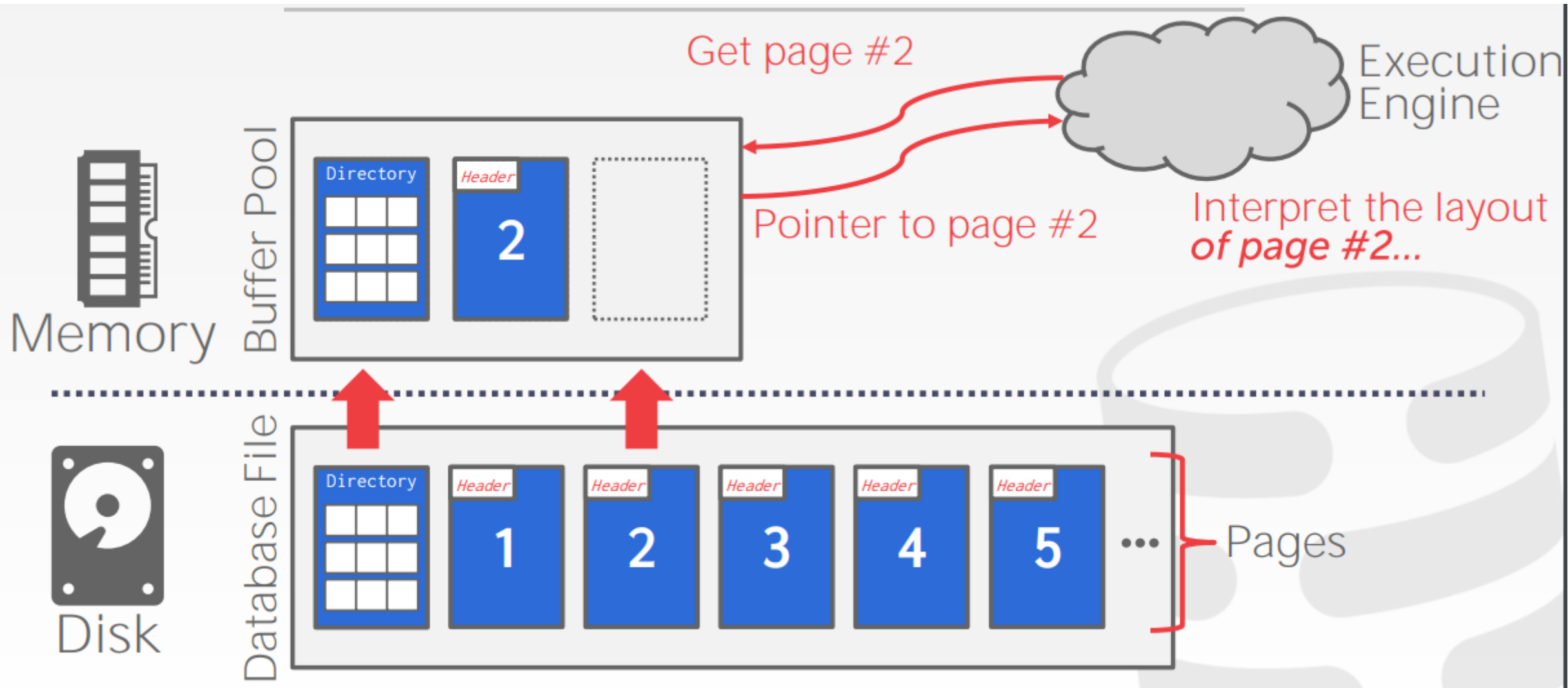


Upravljanje prostorom na disku

Disk space management / Storage manager komponenta



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
-

Fajl sistem i DBMS

- DBMS skladišti bazu podataka u jednom ili više fajlova, pri čemu sam odlučuje kako će organizovati podatke unutar njih.
 - Svaki DBMS ima sopstvenu filozofiju organizacije baze po fajlovima. (fajlovi različitih DBMS-ova uglavnom nisu kompatibilni, tj. ne koriste univerzalne formate).
 - OS ne zna ništa o sadržaju tih fajlova.
-

Ciljevi dizajna DBMS-a kada su u pitanju skladišta

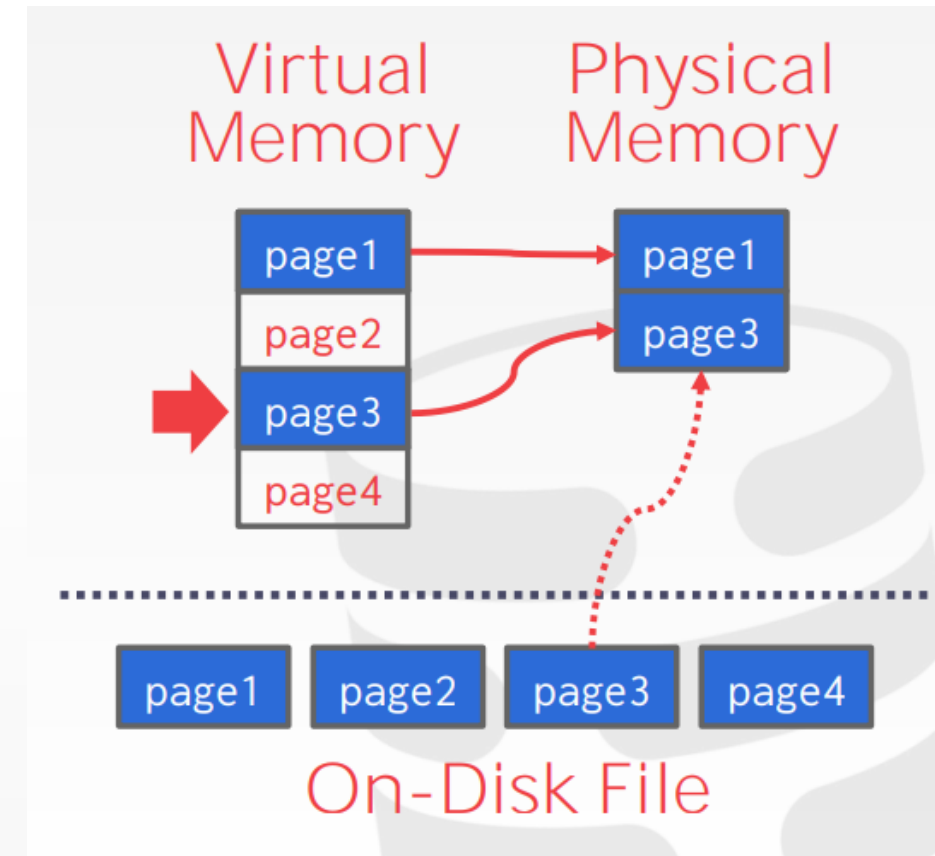
- Omogućiti DBMS-u da upravlja bazama podataka koje prevazilaze kapacitete dostupne memorije -> **trajna skladišta (disk)**.
- Čitanje/pisanje na disk su skupe operacije -> **optimizacija broja takvih operacija**.
- Direktan pristup podacima na disku je značajno sporiji od sekvencijalnog čitanja -> **obezbediti veći procenat ekvencijalnih 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 smeštaju podatke u susedne blokove.

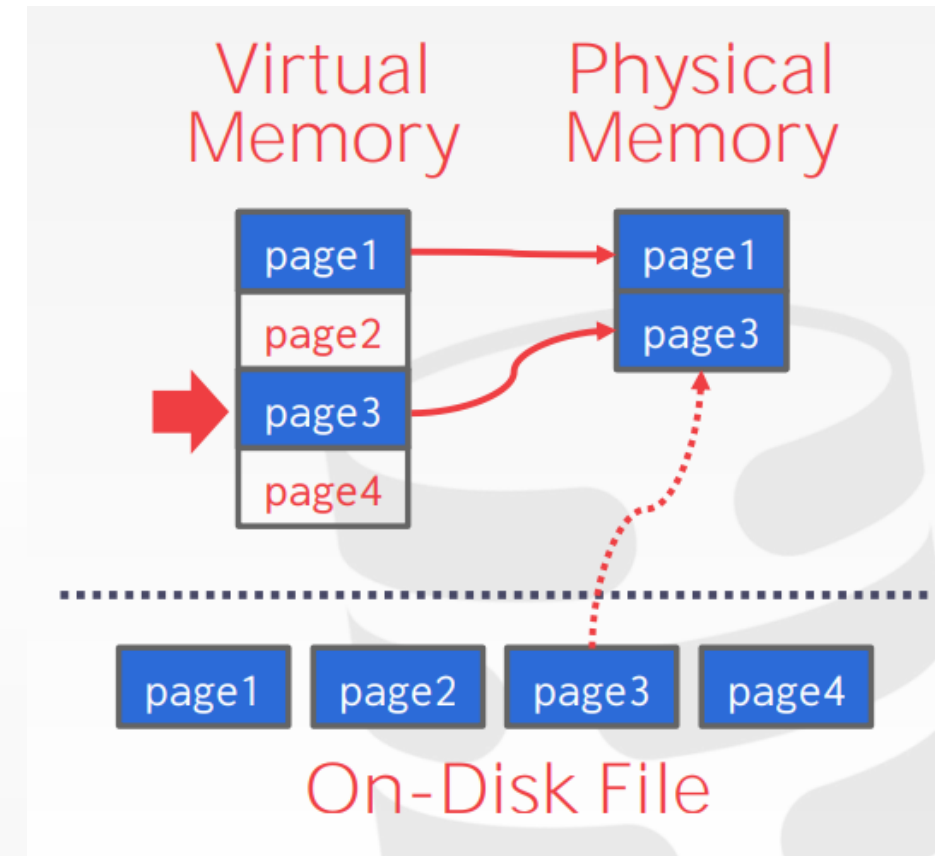
Da li koristiti OS za upravljanje učitavanjem?

- Može se koristiti mapiranje. OS je odgovoran za premeštanje podataka, tj. za pomeranje stranica/blokova fajlova u i iz radne memorije.
- Problem 1: Konkurentan pristup više transakcija nije kontrolisan. OS ne zna ništa o tome da li je neka transakcija završena i može da ukloni stranicu iz virtuelne memorije pre nego što se potvrdi da je transakcija završena



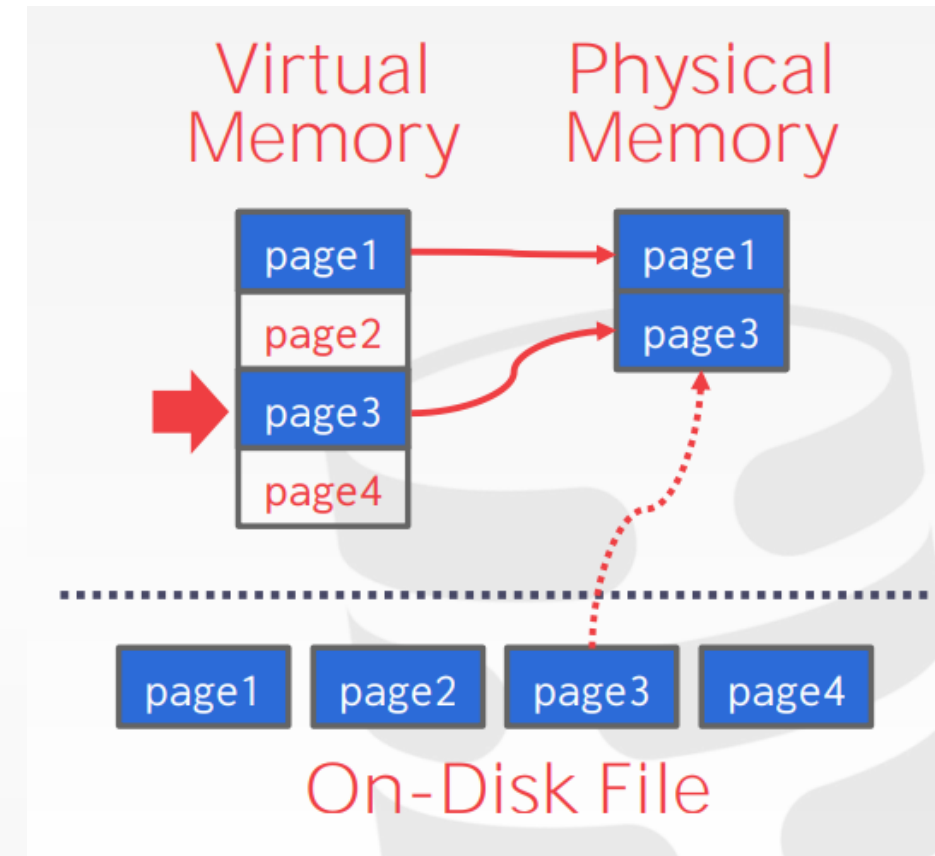
Da li koristiti OS za upravljanje učitavanjem?

- Problem 2: Kašnjenje I/O operacija
DBMS nema uvid u to koje su stranice u virtuelnoj memoriji, jer time upravlja OS: Kada pokuša da pristupi stranici koja nije u memoriji može da dođe do page fault statusa, kada nit DBMS-a koja je tražila stranicu mora da bude na čekanju.
- Problem 3: Upravljanje greškama
Nema lakog načina da DBMS proveri validnost stranice pre pristupa, jer OS kontroliše pristup. DBMS mora sam da presretne i obradi SIGBUS, koji može nastati, na primer. ako DBMS pokuša da pročita deo fajla koji više ne postoji (npr. fajl je kraći nego što se očekuje, ili je oštećen).



Da li koristiti OS za upravljanje učitavanjem?

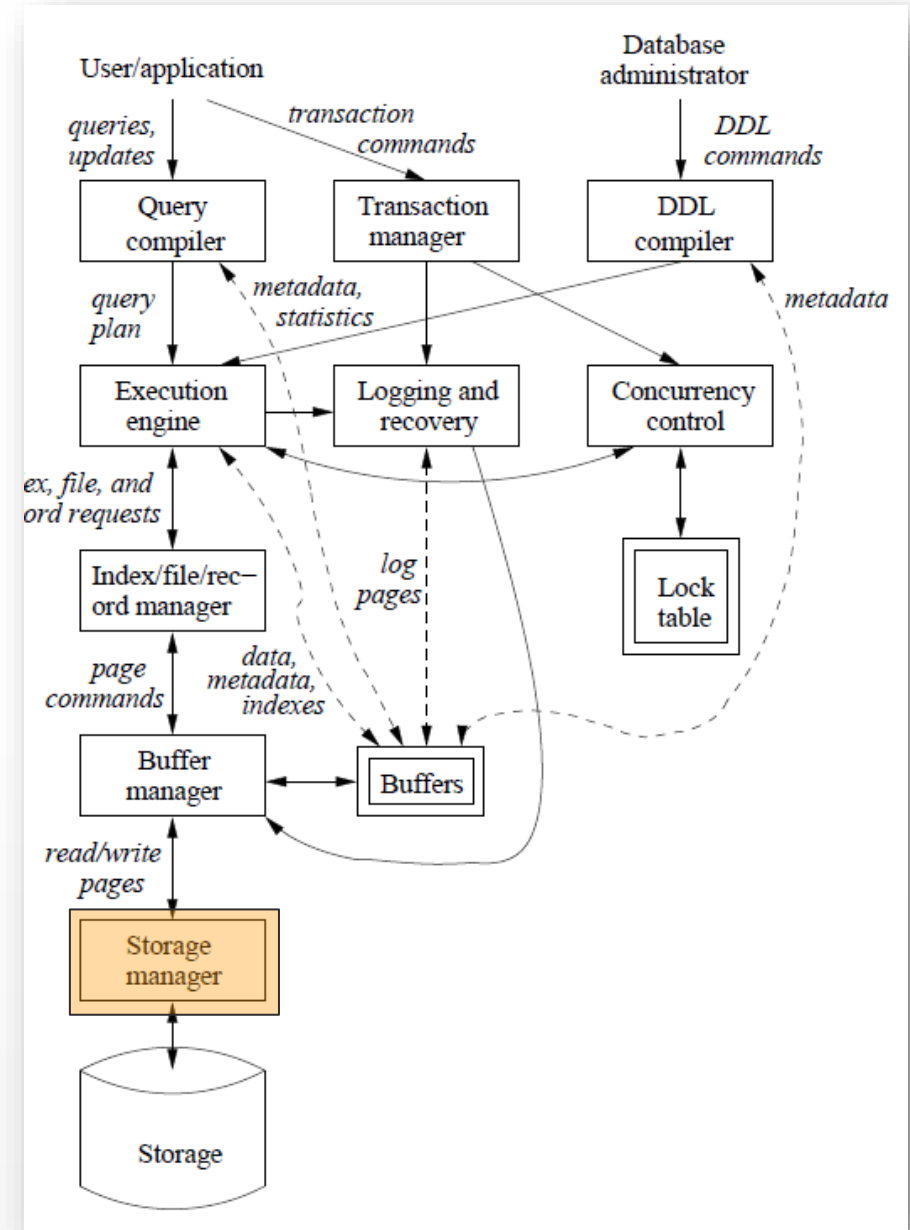
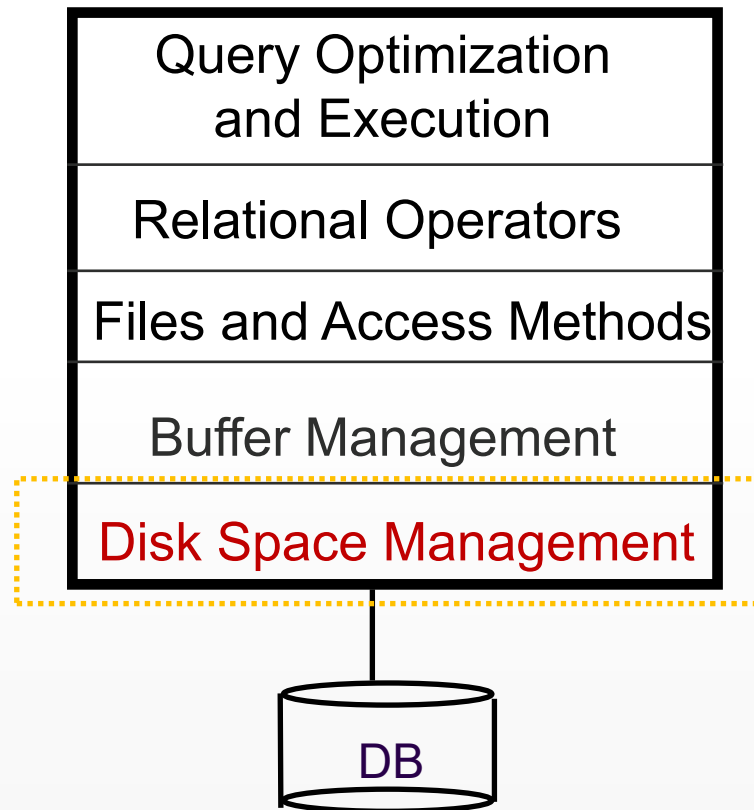
- Problem 4: Loše performance
Potencijalno veliki broj proces konkurentno pristupa metastrukturama virtualne tabele što obara performanse procesa (zbog zaključavanja)
- Može se rešiti, ali je skupo.



Šta je najčešće DBMS-u potrebno?

- Slanje (flushing) izmenjenih (prljavih) stranica na disk u ispravnom redosledu.
 - Odgovarajuće preuzimanje stranica pre nego što su potrebne (pre-fetching).
 - Politika zamene strana.
 - Raspored niti/procesa.
-

DBMS kontekst

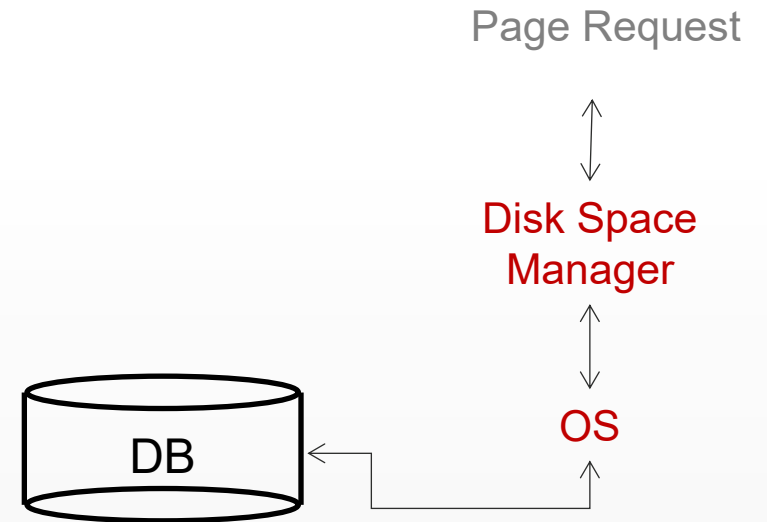


Upravljanje skladištenjem podataka

- 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 DMBSovi u pitanju
-

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.



Storage manager

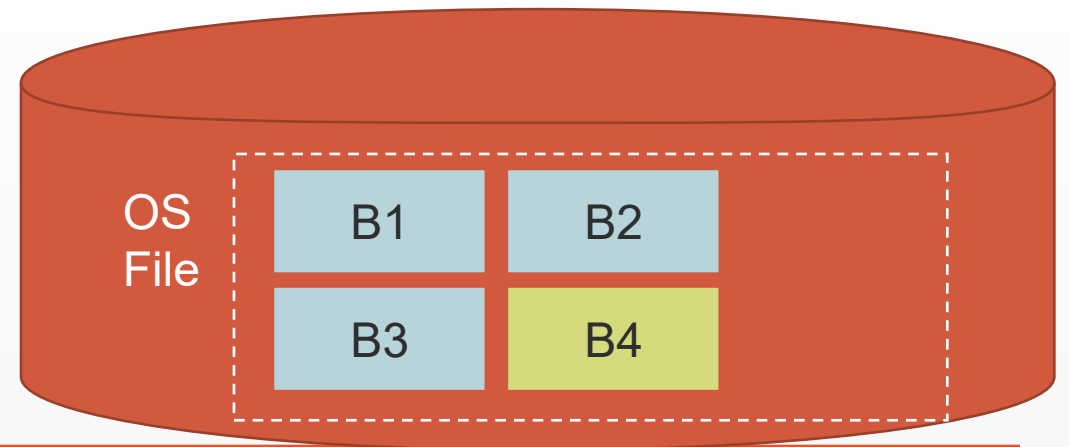
- **Storage Manager (menadžer skladišta/prostora)** – najniži nivo u DBMS arhitekturi
 - Uloga menadžera prostora jeste upravljanje prostorom na disku.
 - DBMS čuva bazu podataka u jednom ili više fajlova (OS).
 - Jedan OS fajl na fajl sistemu se sastoji iz više OS blokova, ali ih kao takve u DBMS-u vidi samo Menadžer skladišta.
-

Storage manager

- Podaci se sa diska dopremaju u radnu memoriju u OS blokovima.

Veličina bloka – od 512 bajtova do nekoliko kilobajta

- Manji blokovi – više transfera
- Veći blokovi – više protraćenog prostora

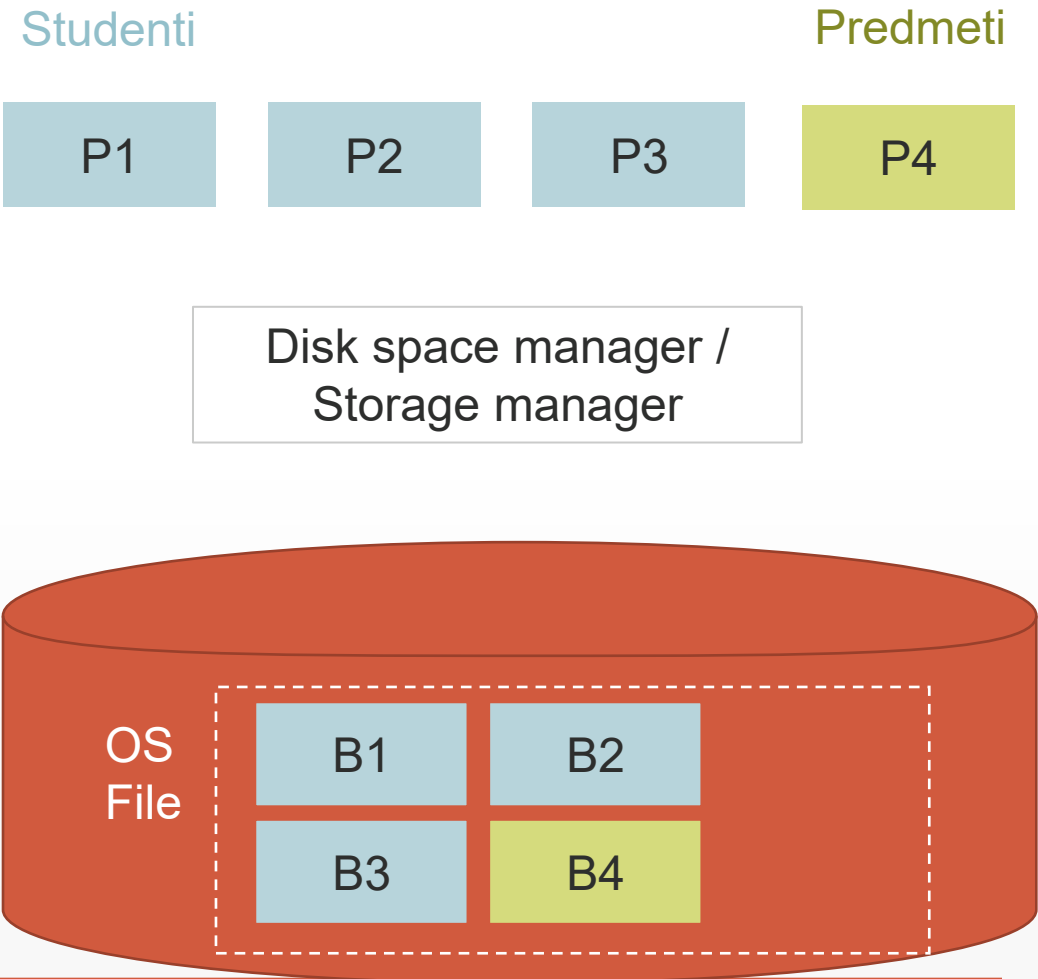


Storage manager

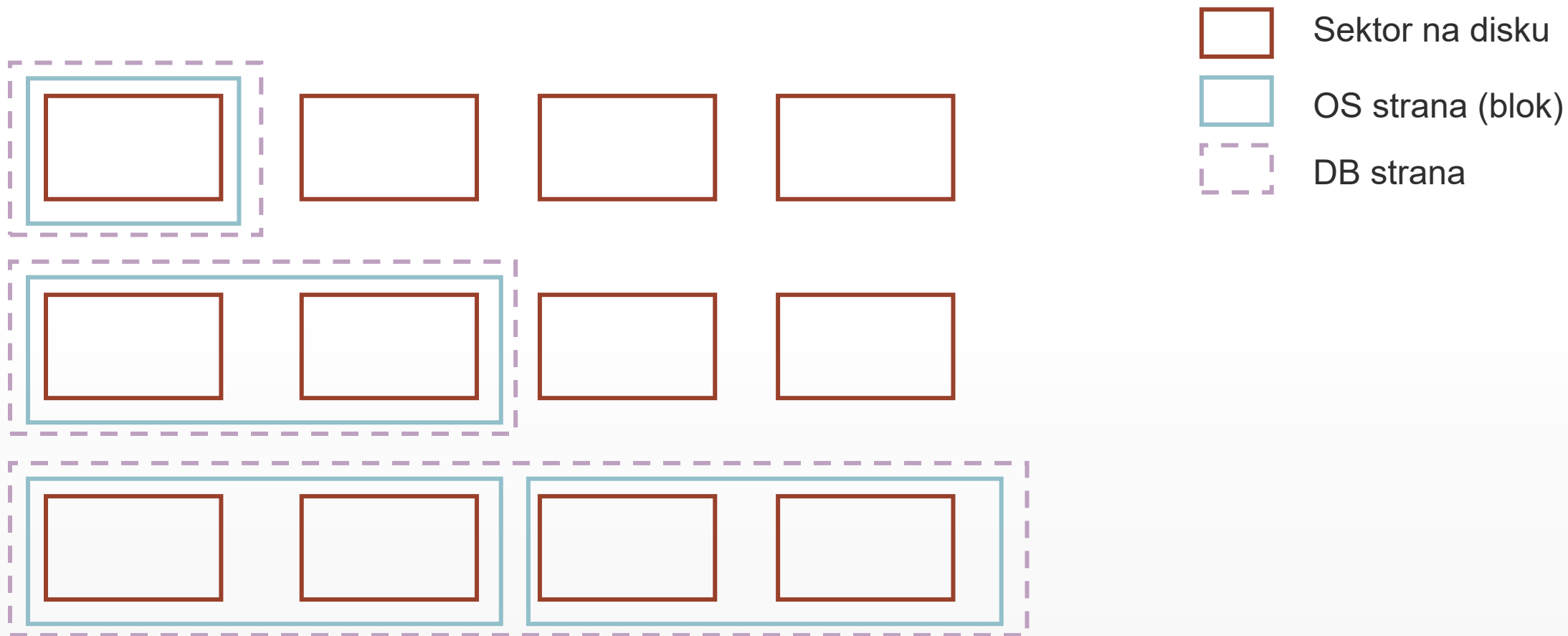
- Menadžer skladišta podržava **koncept strane (page)** kao jedinice podataka i **obezbeđuje** komande za **alociranje i oslobađanje**, kao i **čitanje i pisanje u stranu**.

the size of the page = $k * \text{size of a disk block}$

- Veličina strane je fiksna.
- Svaka DB strana ima jedinstven identifikator.
Menadžer skladišta mapira identifikatore strana na fizičke lokacije na disku (OS blokove).



Blokovi i strane



Veličina stranica

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

- Hardware Page (usually 4KB)
- OS Page (usually 4KB)
- Database Page (512B-16KB)

A hardware page is the largest block of data that the storage device can guarantee failsafe writes.

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.
-