

Skladišta podataka

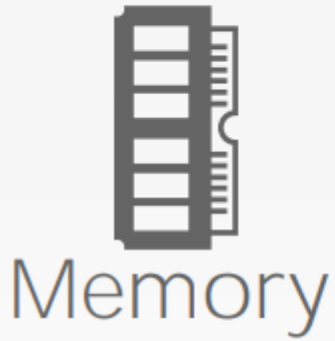
STORAGES

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**.

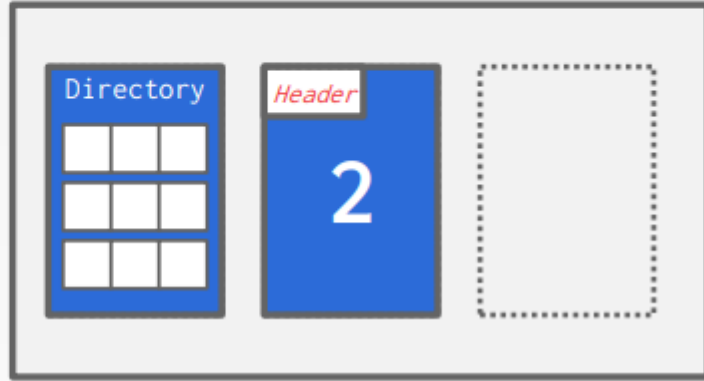
Tradicionalni DBMS-ovi su dizajnirani da maksimizuju zastupljenost sekvencijalnog čitanja.

Algoritmi za optimizaciju teže da smanje broj direktnih pristupa stranicama tako što se trude da speštaju podatke u susedne blokove.



Memory

Buffer Pool



Get page #2

Pointer to page #2



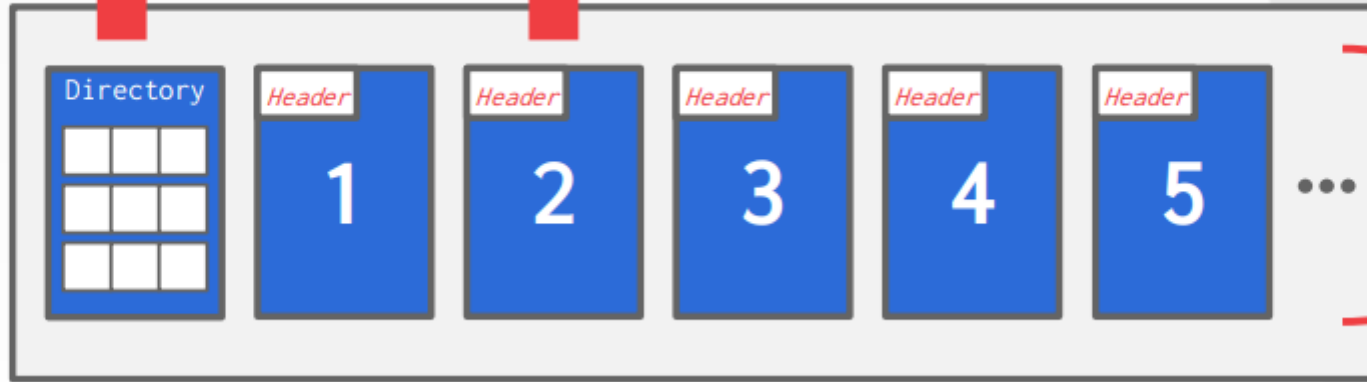
Execution Engine

Interpret the layout of page #2...



Disk

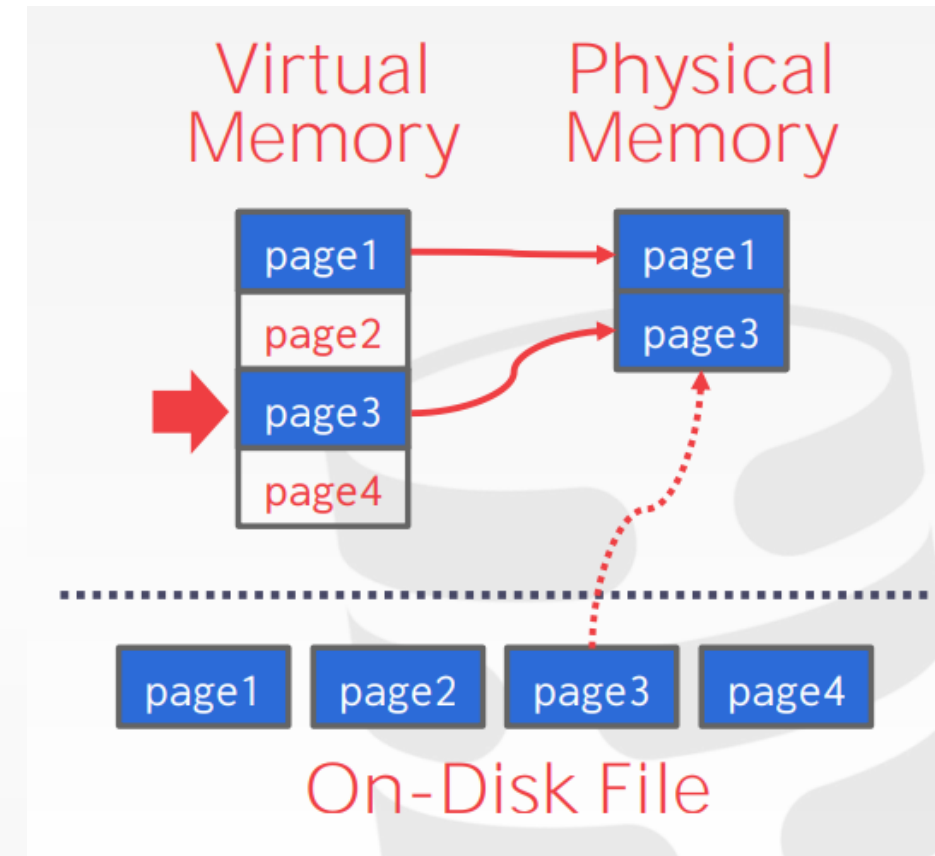
Database File



Pages

Da li koristiti OS za upravljanje učitavanjem?

- Može se koristiti mapiranje.OS je odgovoran za premeštanje podataka za pomeranje stranica datoteka u i iz radne memorije.
- Moguće je:
 - OS kako očekujete da ćete čitati određene stranice
 - Zabraniti uklanjanje određenih stranica iz memorije (jer se koriste)
 - Tražiti da se određene stranice uklone iz memorije (jer se više ne koriste).



Da li koristiti OS za upravljanje učitavanjem?

- Može se koristiti mapiranje. OS je odgovoran za premeštanje podataka za pomeranje stranica datoteka u i iz radne memorije.
- Moguće je:
 - OS kako očekujete da ćete čitati određene stranice
 - Zabraniti uklanjanje određenih stranica iz memorije (jer se koriste)
 - Tražiti da se određene stranice uklone iz memorije (jer se više ne koriste).



Šta je najčešće DBMS-u potrebno?

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

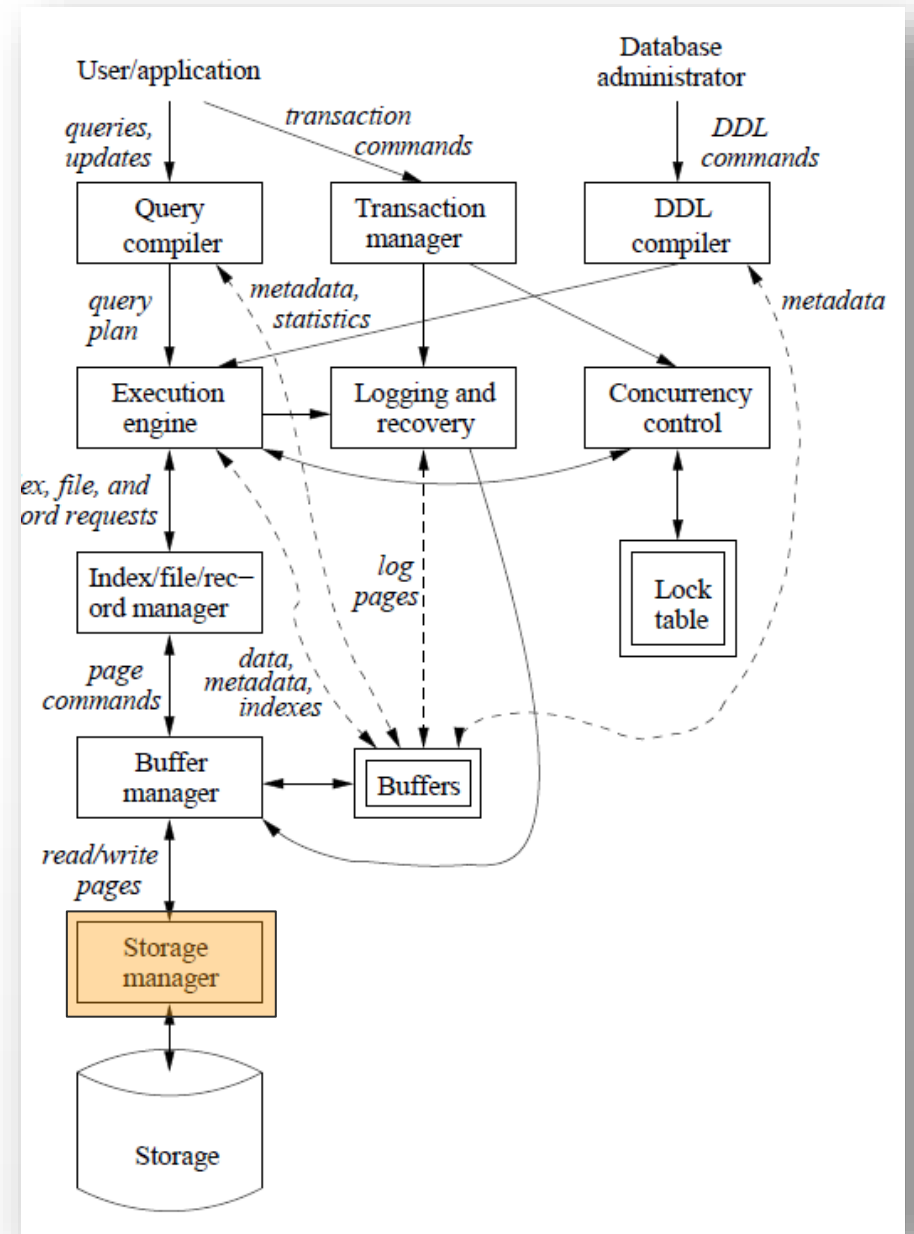
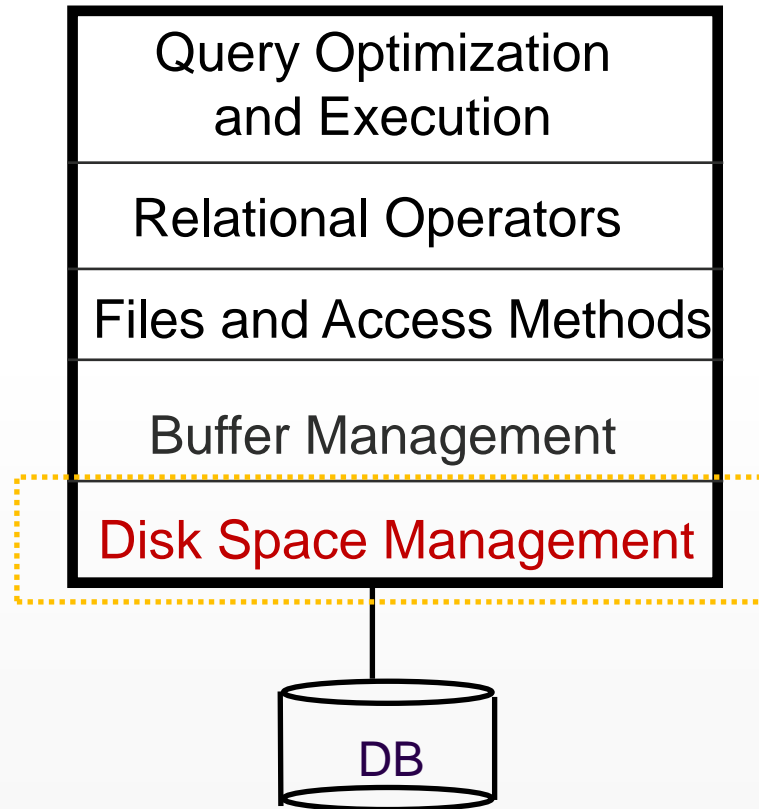
Upravljanje prostorom na disku

Disk space management / Storage manager komponenta

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 saraduje 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
-

DBMS kontekst



Storage manager

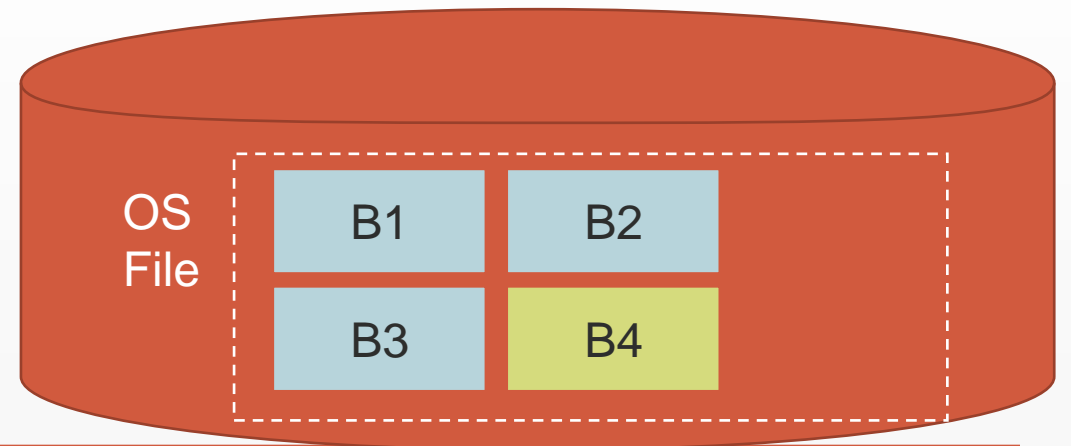
- **Storage Manager (menadžer skladišta/prostora)** – najniži nivo u DBMS arhitekturi
 - Uloga menadžera prostora jeste upravljanje prostorom na disku.
 - DBMS čuva bazu podataka u jednom ili više fajlova (OS).
 - Jedan fajl na fajl sistemu se sastoji iz više OS blokova, ali ih kao takve u DBMS-u vidi samo Menadžer skladišta.
 - Svaki DBMS ima sopstvenu filozofiju organizacije baze po fajlovima.
 - OS ne zna ništa o sadržaju tih fajlova.
-

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

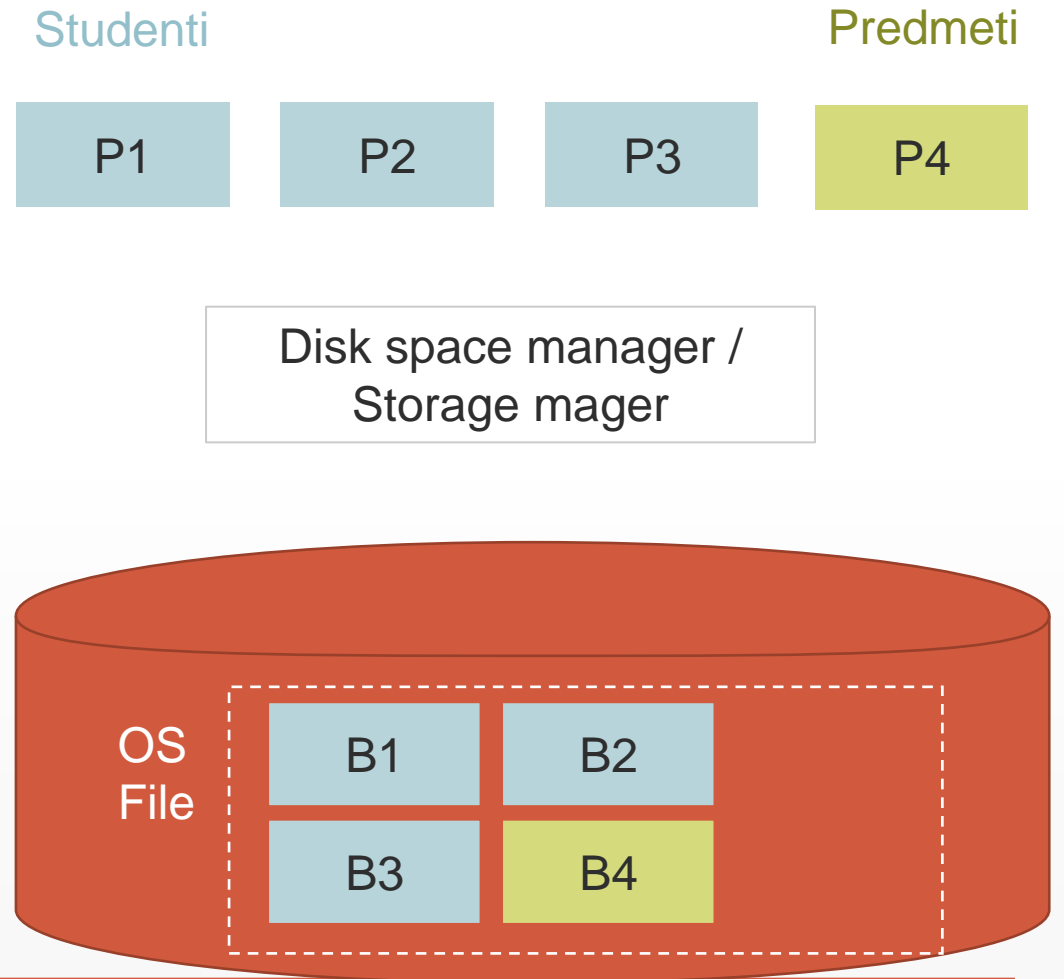


Storage manager

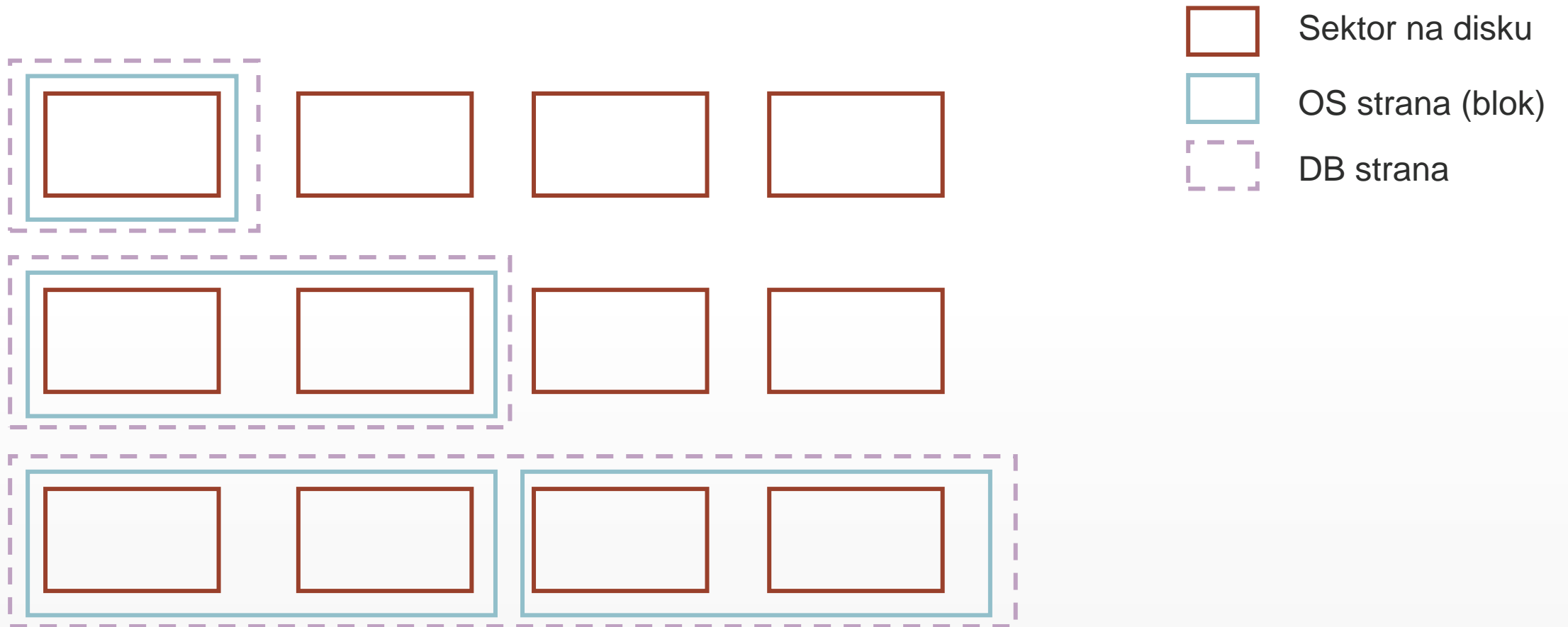
- Menadžer prostora podržava **koncept strane (page)** kao jedinice podataka i **obezbeđuje** komande za **alociranje i oslobađanje**, kao i **čitanje i pisanje u stranu**.

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

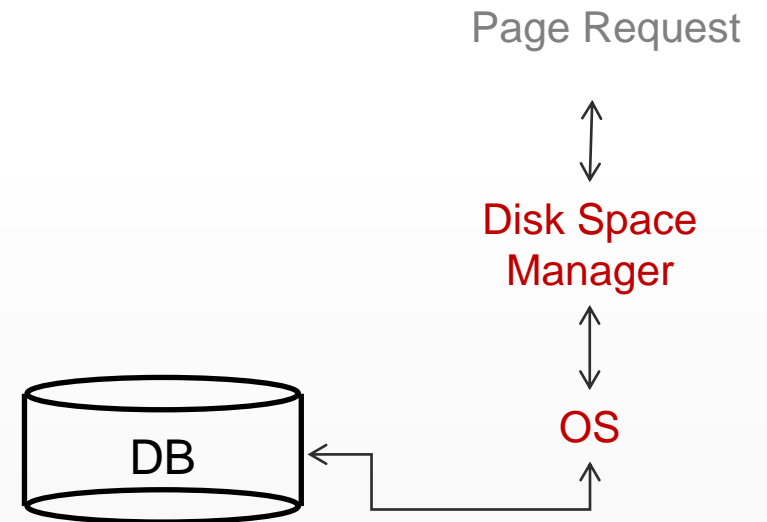


Upravljanje prostorom na disku

- 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.
-

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.

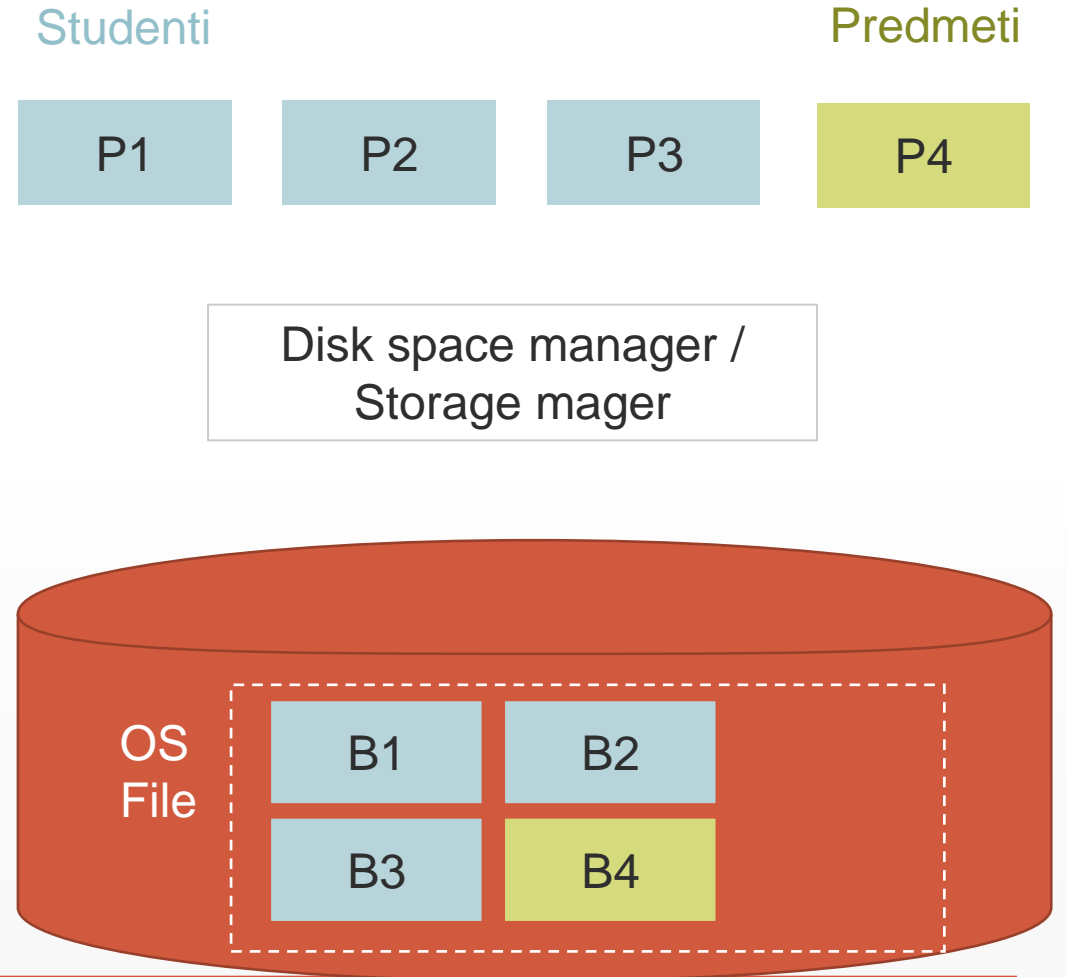


DBMS 'virtuelni' fajlovi

File of records

File of records

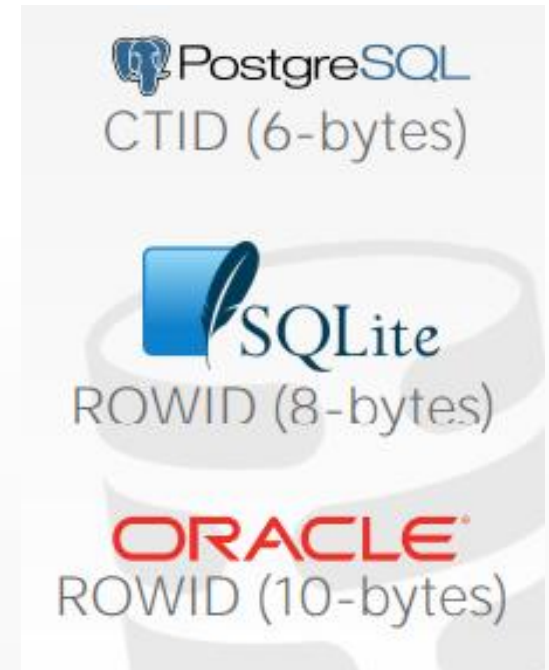
- Sve komponente koje rade sa podacima ih posmatraju u kontekstu kolekcija slogova, tj. **slogovnih fajlova** (*file of records*).
- Slogovni fajl obično sadrži slogove jedne tabele.
- Slog – n-torka polja (*fields*).
- Slogovni fajlovi su logičke organizacione jedinice i predstavljaju kolekciju strana.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



Slogovni fajlovi

- Menadžer fajlova (index/file/record manager) formira i održava fajl sa podacima 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/offset)

Koji se naziva identifikatorom sloga, **record id (RID)**

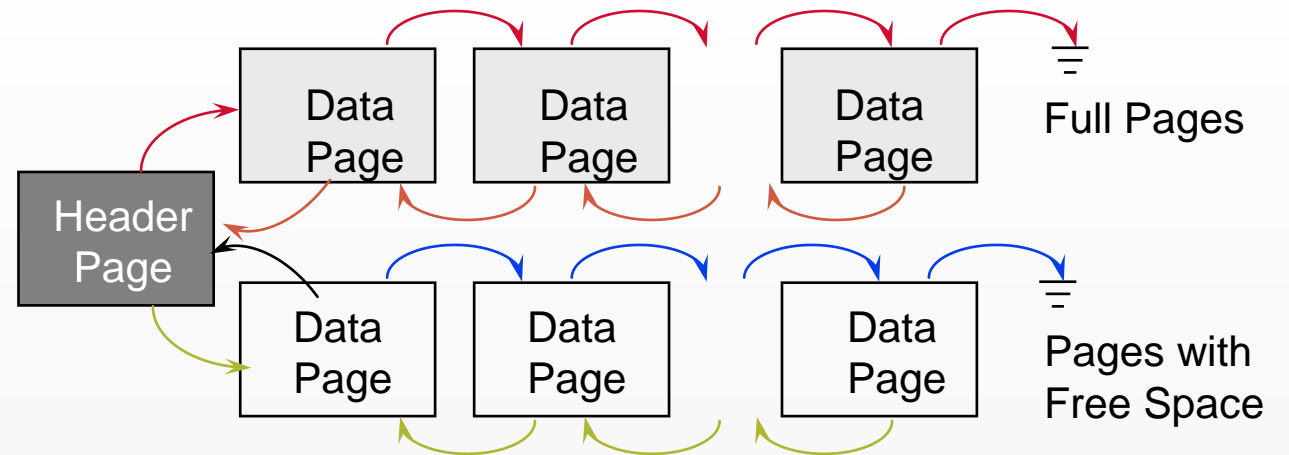


Neuređeni (heap) fajlovi

- Najjednostavnija DB fajl struktura koja ne podrazumeva nikakvo uređivanje slogova unutar fajla.
 - 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
 - Treba voditi evidenciju o:
 - Slobodnom prostoru unutar strane i
 - Stranama koje imaju slobodnog prostora
-

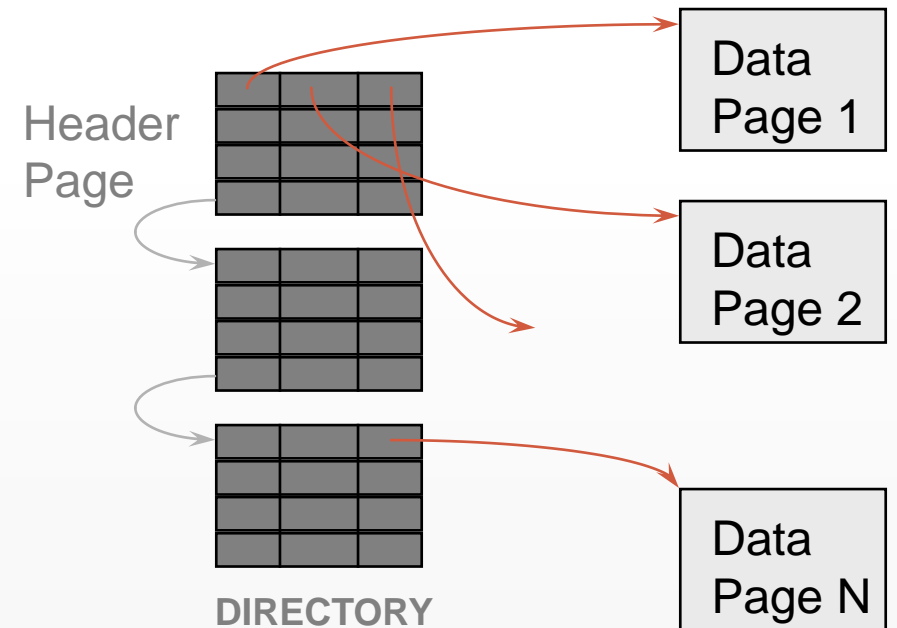
Heap fajl implementiran kao lista

- 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)
- Svaka strana čuva informaciju o slobodnim slotovima.



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



Organizacija podataka na strani

Page layout

Postavka strane (page layout)

- DB fajl može sadržati:
 - Podatke (iz tabela) – slogovni fajlovi
 - Metapodatke
 - Indekse
 - Logove
-

Postavka strane (page layout)

- Svaka strana sadrži zaglavlje (header) u kom se nalaze metapodaci o sadržaju strane:
 - Veličina strane
 - Kontrolnu sumu (checksum)
 - Verziju DBMSa
 - Informacije o kompresiji
 - Vidljivost od strane transakcija
-

Postavka strane (page layout)

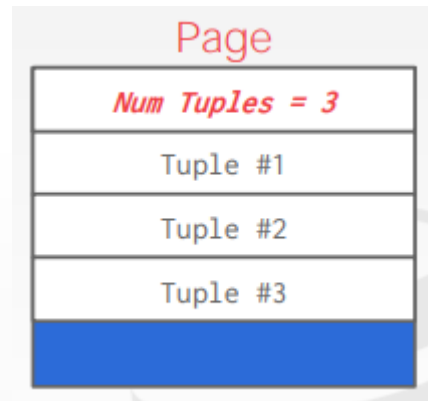
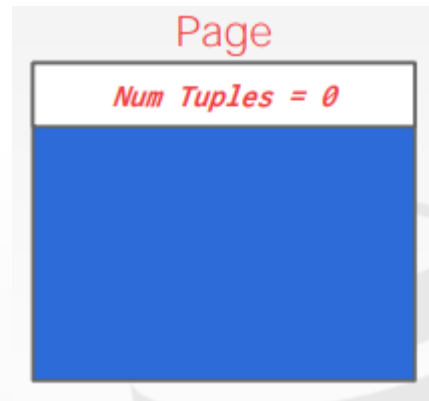
- Kao i
 - File ID
 - Page ID
 - Tip strane (npr. strana prekoračenja)
 - Vremeski pečat
 - Tabela/directory slogova
 - Pointer na slobodan prostor
-

Dve vrste u organizaciji postavke

- Dva osnovna pristupa u zavisnosti od sadržaja:
 - Prilagođen pamćenju torki (tuple-oriented)
Potencijalno raličita rešenja kada su u pitanju:
 - Slogovi fiksne dužine
 - Slogovi varijabilne dužine
 - Prilagođen beleženju log podataka (log-structured)
-

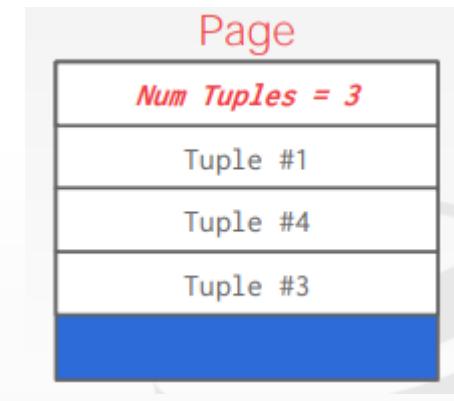
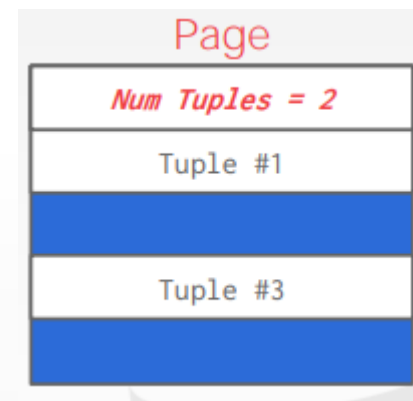
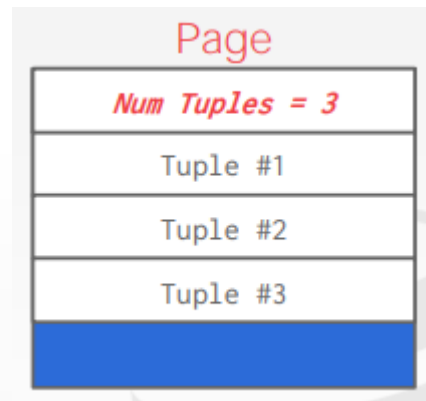
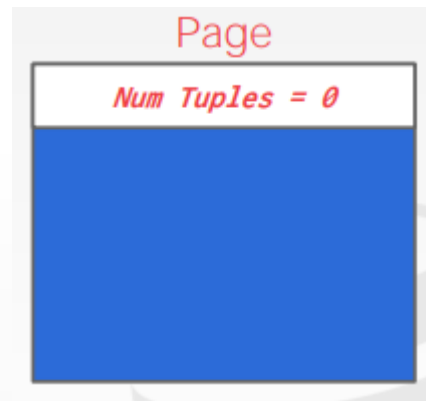
“Torka” strane sa slogovima fiksne dužine

- Jedan način: Pamtiti broj zauzetih mesta.

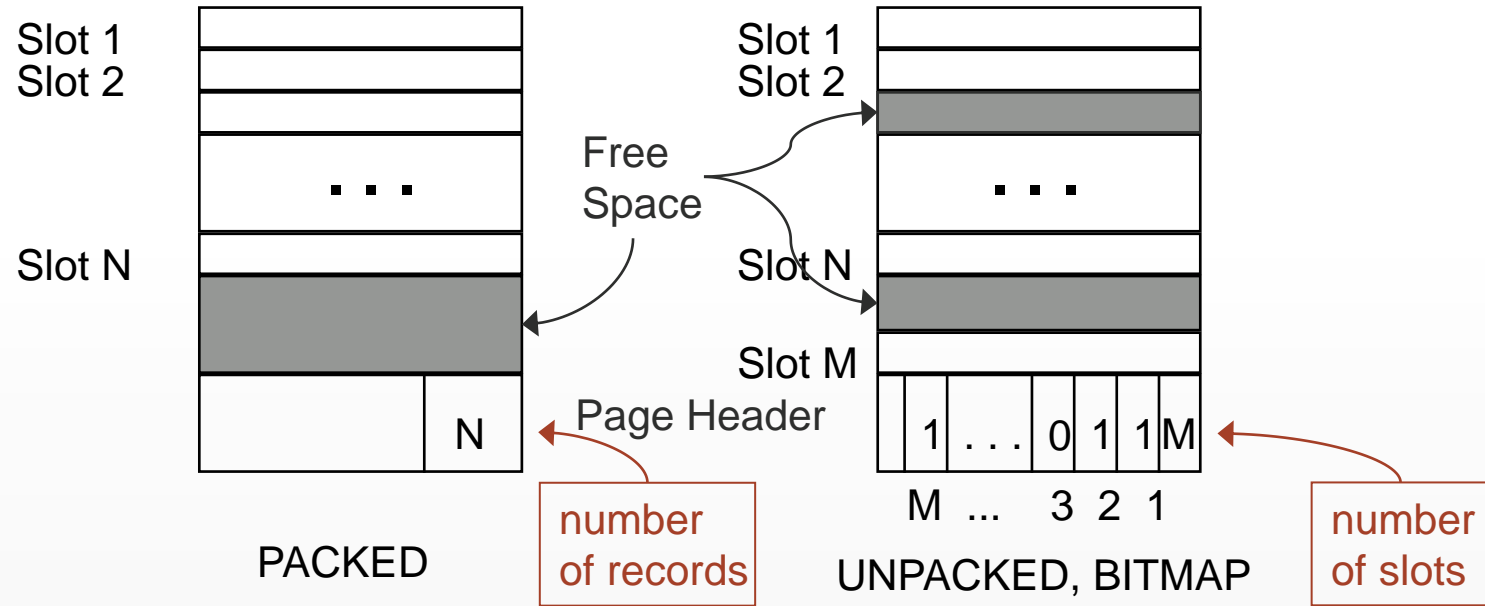


“Torka” strane sa slogovima fiksne dužine

- Šta sa brisanjem?

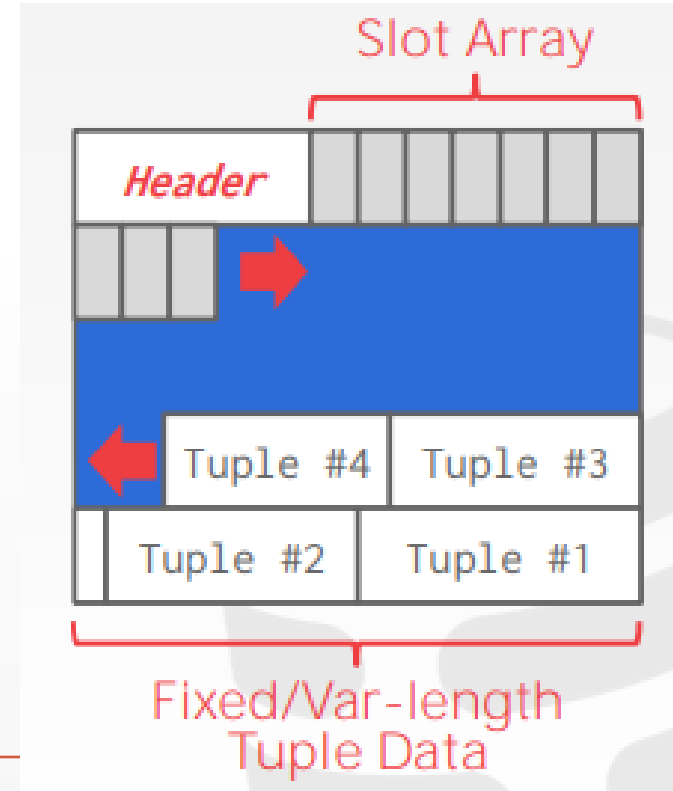


“Torka” strane sa slogovima fiksne dužine



“Torka” strane sa slogovima fiksne/varijabilne dužine

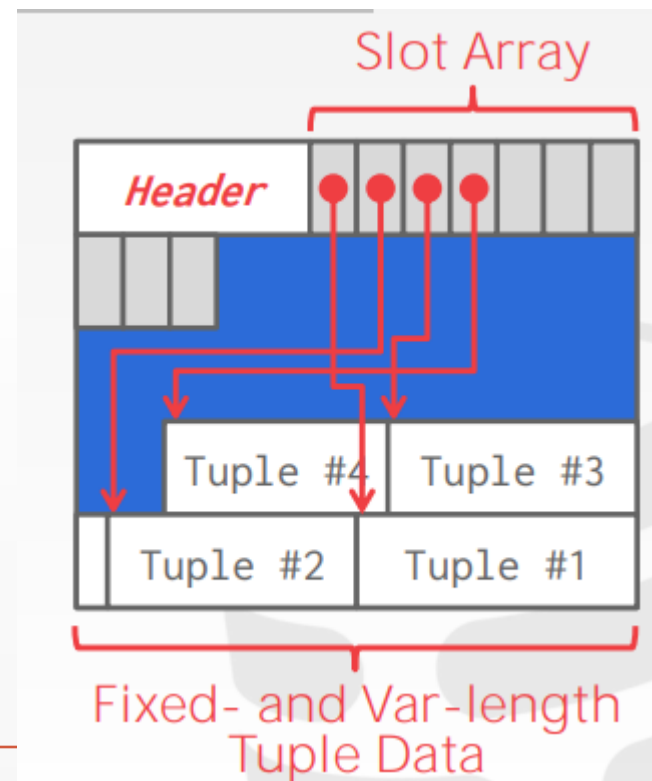
- Najfleksibilniji način organizacije strane sa slogovima i fiksne i promenljive dužine je održavanje tabele/mape slotova
- Najšešća postavka. Naziv: *slotted pages*.



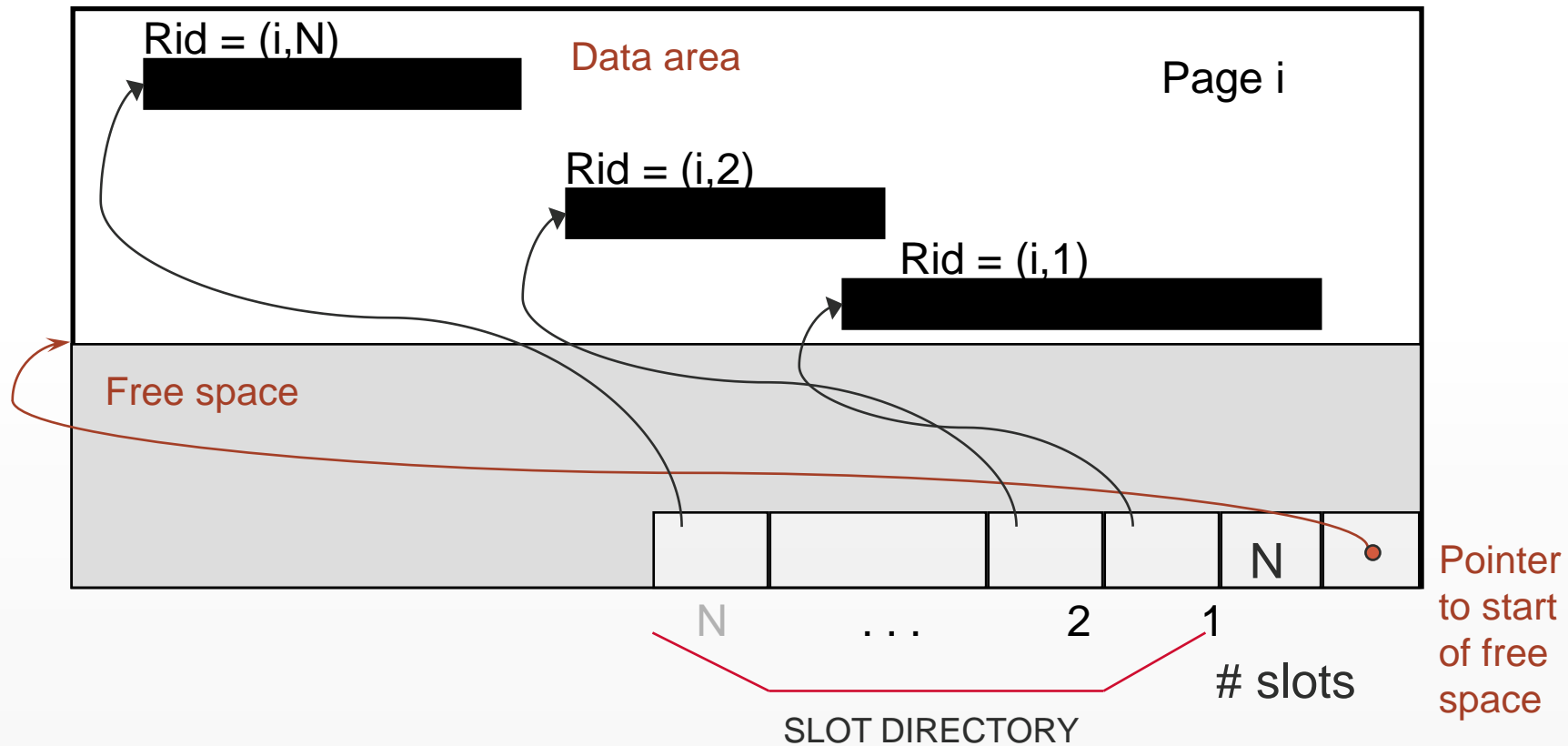
“Torka” strane sa slogovima fiksne/varijabilne dužine

- Svaki slot opisuje parom
(record offset, record length)
- Offset – broj bajtova od početka prostora sa podacima.
- Brisanje – prepakovati ili postaviti offset na -1, pa prepakovati kada se ispostavi da nema mesta za nove.
- Slot kojem je pridružen obrisani slog ne može uvek biti obrisan, jer bi poremetio brojanje, tj. redne brojeve slotova, a time potencijalno i rid-ove postojećih slogova.

RID (pageID, slot number/offset)

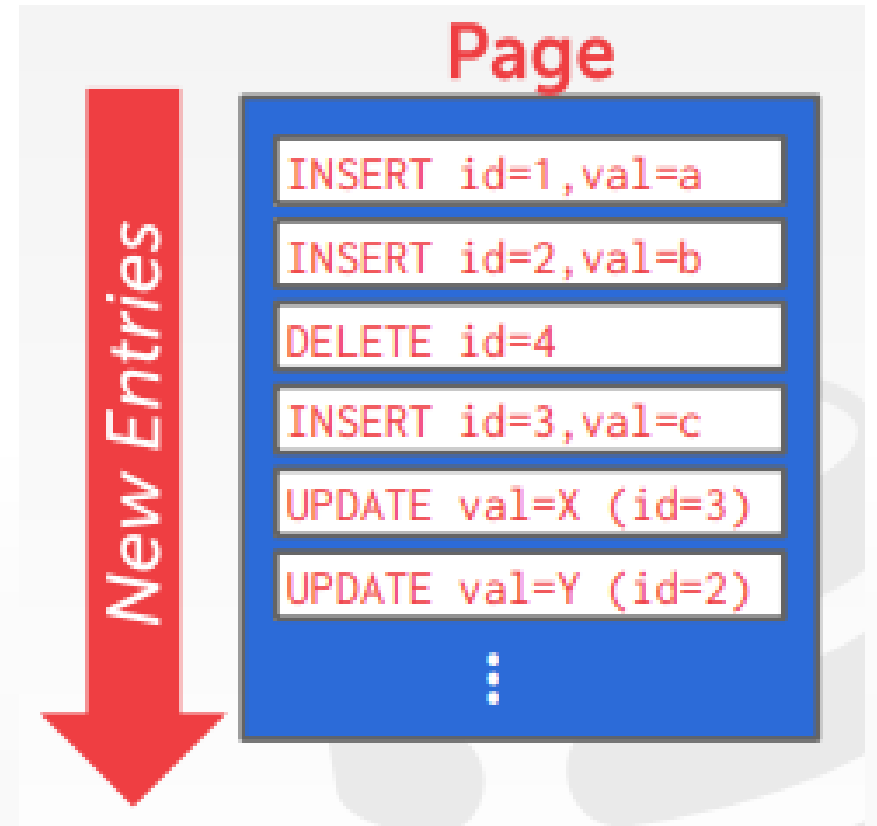


“Torka” strane sa slogovima fiksne/varijabilne dužine



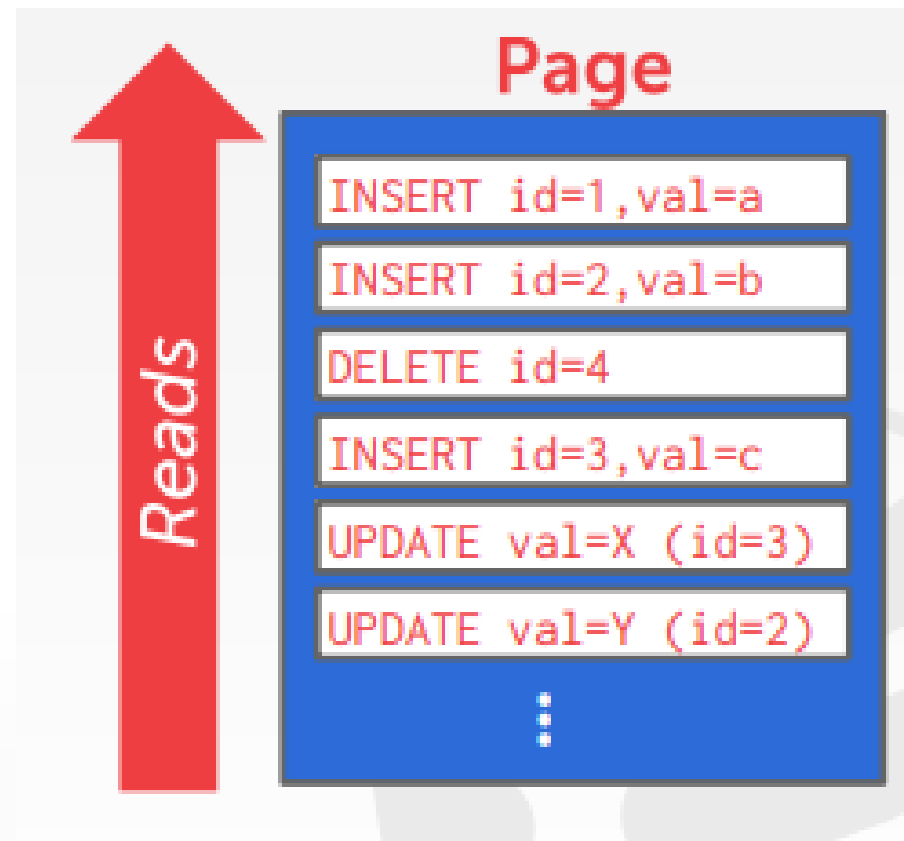
Log strane

- Beleženje zapisa o izmenama izvršenim nad bazom podataka (logovi).
- **Kakva je potreba za kojom CRUD komandom?**



Log strane

- Beleženje zapisa o izmenama izvršenim nad bazom podataka (logovi).
- Češće pisanje nego čitanje.
- Pretraga sekvencijalna i unazad.



APACHE
HBASE



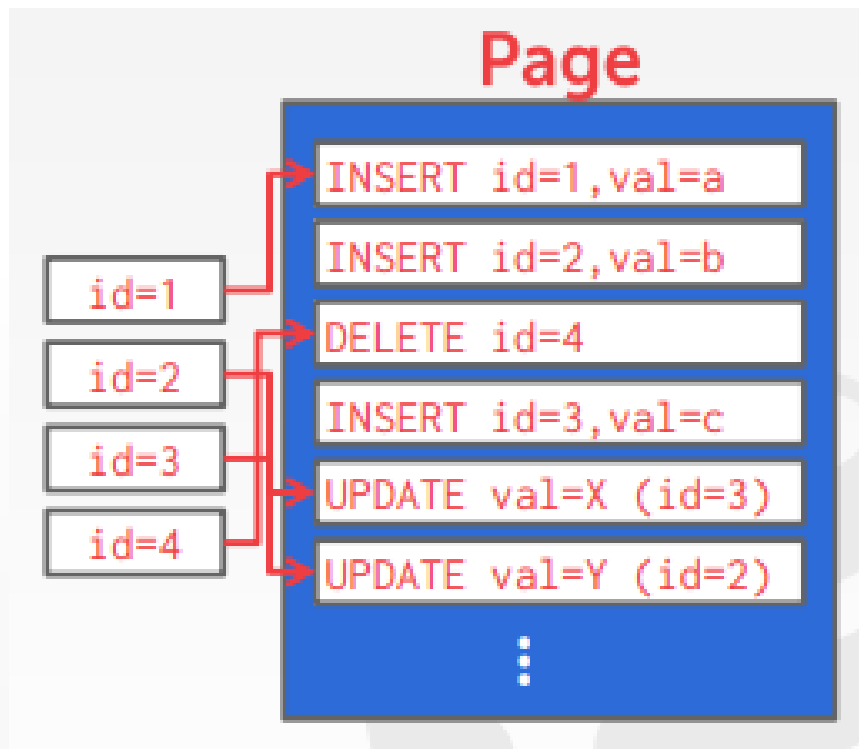
levelDB



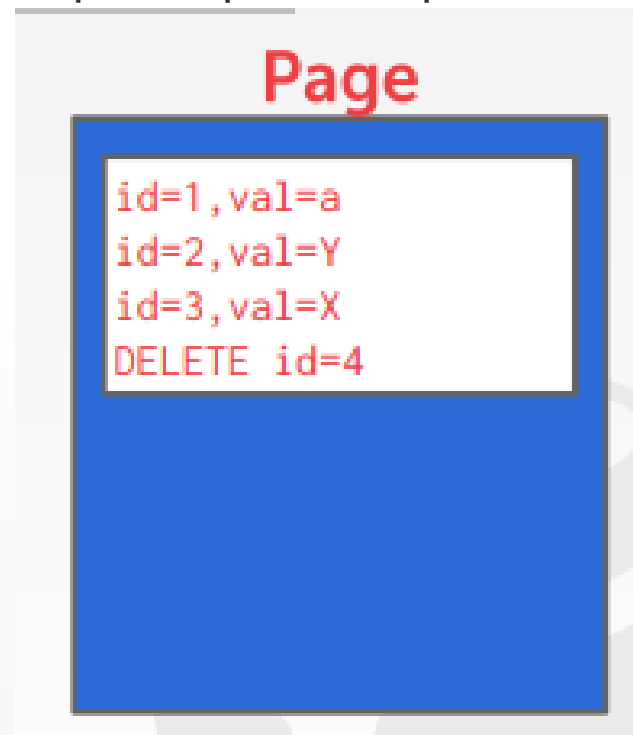
RocksDB

Log strane

- Moguće postavljanje indeksa.



- Periodična kompresija uklanjanjem nepotrebnih zapisa. Npr. Više update-ova u jedan.



Organizacija podataka u slogu

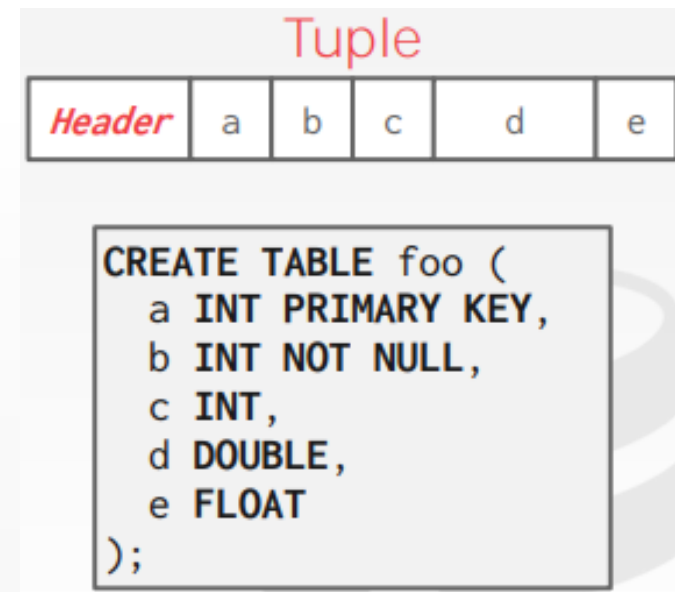
Tuple layout

Slogovi i polja (vrednosti atributa)

- Slog predstavlja kolekciju polja (koja su slogom povezana u logičku celinu)
 - Slog ima definisanu strukturu / format
 - Polje nosi podatak koji je određenog tipa
 - Slogovi se beleže kao niz bitova/bajtova. **Zadatak DBMSa je da ih protumači i izdvoji pojedine delove kao vrednosti atributa jedne torke.**
 - Pitanja:
 - Kako se beleže formati?
 - Kako se beleže pojedini tipovi podataka?
- Dodatno: kako oba utiču na organizaciju slogova u fajlu?
-

Tuple layout

- Slogovi mogu sadržati zaglavlje u kom se mogu naći
 - Tip sloga ili celu šemu sloga
 - Dužina sloga
 - Vremenski pečat
 - ...



Uključivanje šeme u slog

Example: Include record type in record

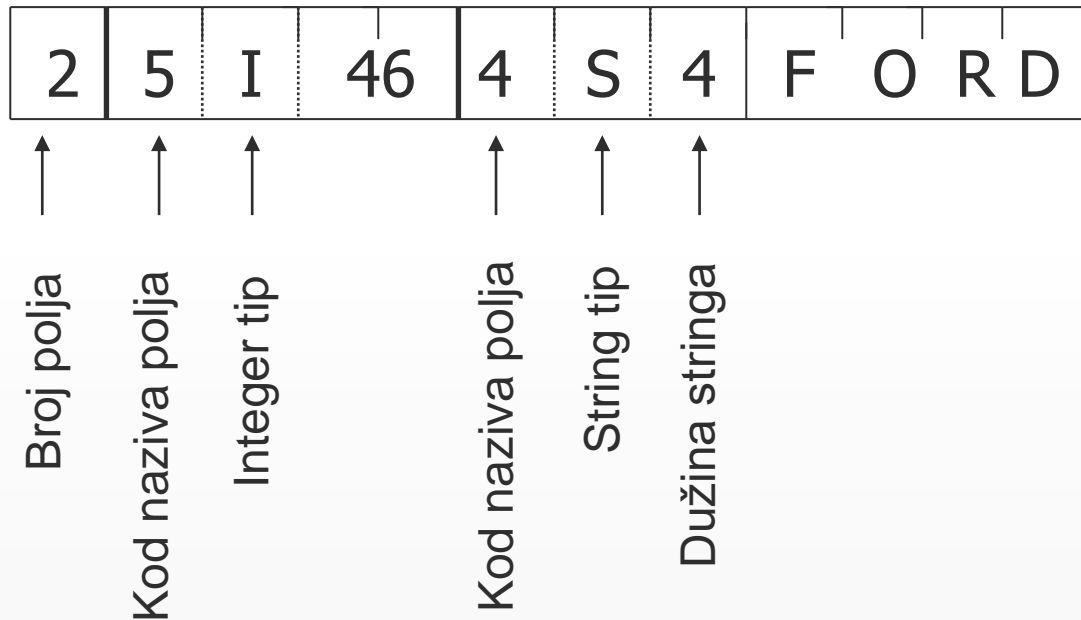


Record type
tells me what
to expect
(i.e. points to
schema)

record length

Uključivanje šeme u slog

- Samoopisujući format



Kodovi naziva polja mogu biti bilo kakvi tagovi.

Tipovi slogova

- FORMAT: FIKSNI ili PROMENLJIVI
- DUŽINA: FIKSNA ili PROMENLJIVA

Struktura, tj. format sloga je opisan **ŠEMOM**.

Tipovi slogova – Fiksni format

- Šema sloga sadrži
 - Broj polja
 - Tip svakog polja
 - Redosled polja
 - Značenje svakog polja

Employee record

(1) E#, 2 byte integer

(2) E.name, 10 char.

(3) Dept, 2 byte code

} Schema

55	s	m	i	t	h							02
----	---	---	---	---	---	--	--	--	--	--	--	----

83	j	o	n	e	s							01
----	---	---	---	---	---	--	--	--	--	--	--	----

} Records

Promenljivi format

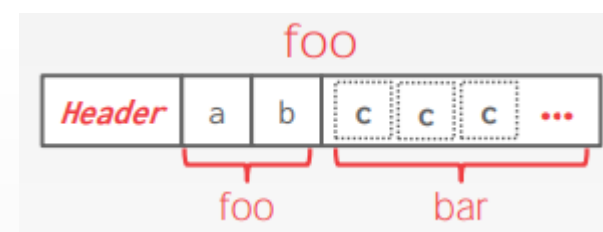
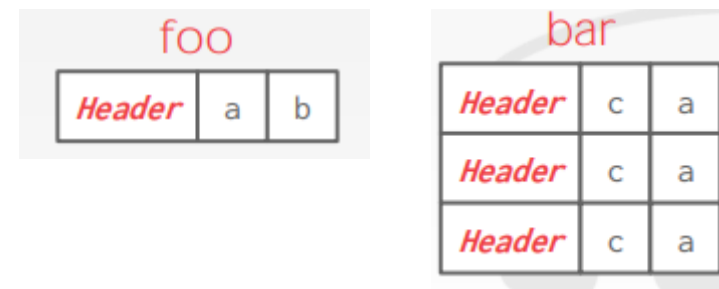
- Pogodan za
 - Retke slogove
 - Ponavljajuća polja
 - Promenljive definice formata
- EXAMPLE: promenljivi format sa ponavljajućim poljima

Employee → one or more → children

3	E_name: Fred	Child: Sally	Child: Tom
---	--------------	--------------	------------

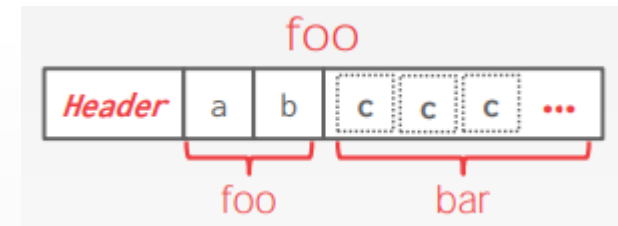
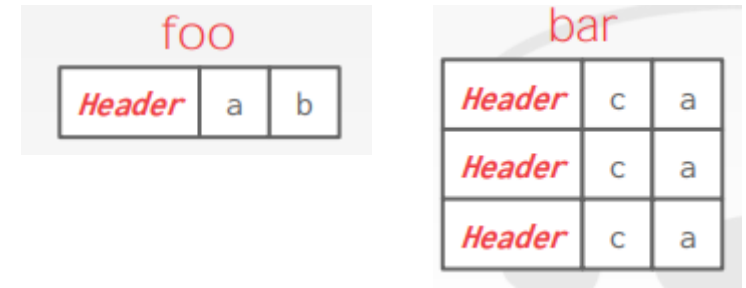
Denormalizovane torke

- Na jednoj strani se mogu naći i torke više tabela, vezanih stranim ključem.
- Prednosti i mane?



Denormalizovane torke

- Na jednoj strani se mogu naći i torke više tabela, vezanih stranim ključem.
- Zahtevi za odgovarajućim join-om će imati manje čitanja.
- Ažuriranje je skuplje.
- Nisu sve pretrage 'jeftinije'.



Četiri osnovna načina beleženja više slogova unutar strane

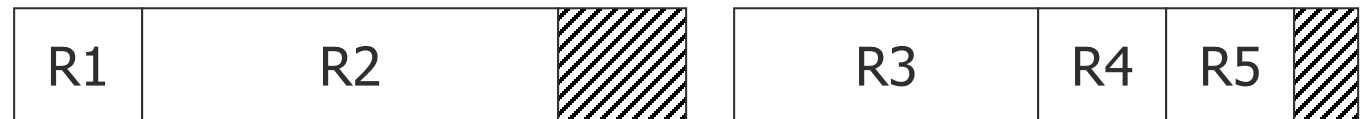
- Razdvajanje slogova markerima



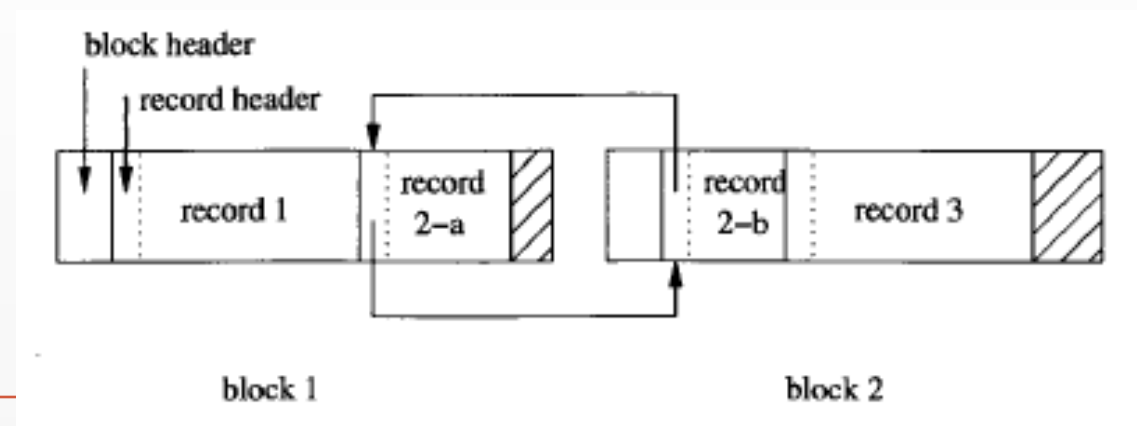
- Slogovi fiksne dužine – ne moraju imati separator
 - Uvođenje specijalnog markera kao separatora
 - Zadavanje dužine sloga (ili ofseta)
 - Unutar svakog sloga
 - U zaglavlju bloka
-

Četiri osnovna načina beleženja više slogova unutar strane

- Razdvajanje slogova markerima
- Razapeti vs. celoviti
- Nerazapeti slogovi – jednostavniji, postojanje neiskorišćenog prostora

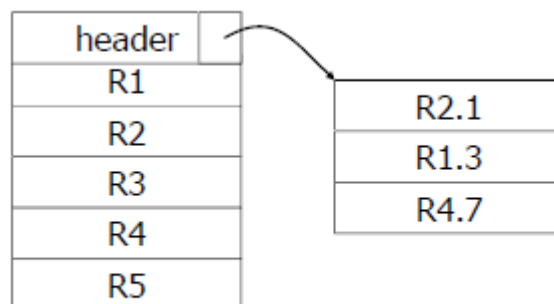
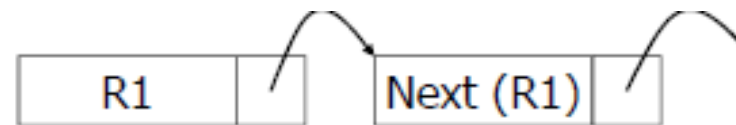


- Razapeti – neophodni kada je veličina sloga veća od veličine sloga



Četiri osnovna načina beleženja više slogova unutar strane

- Razdvajanje slogova markerima
- Razapeti vs. celoviti
- Uređivanje
- Fajlovi čiji su slogovi uređeni prema vrednosti ključa – fajlovi sa sekvencijalnim uređenjem. (uređenje ubrzava pretragu, doduše po ključu)



- Opcije za uređivanje
 - Slogovi su fizički uzastopno smešteni
 - Slogovi čine povezanu listu
 - Prostor prekoračenja

Četiri osnovna načina beleženja više slogova unutar strane

- Razdvajanje slogova markerima
- Razapeti vs. celoviti
- Uređivanje
- Preusmeravanje

