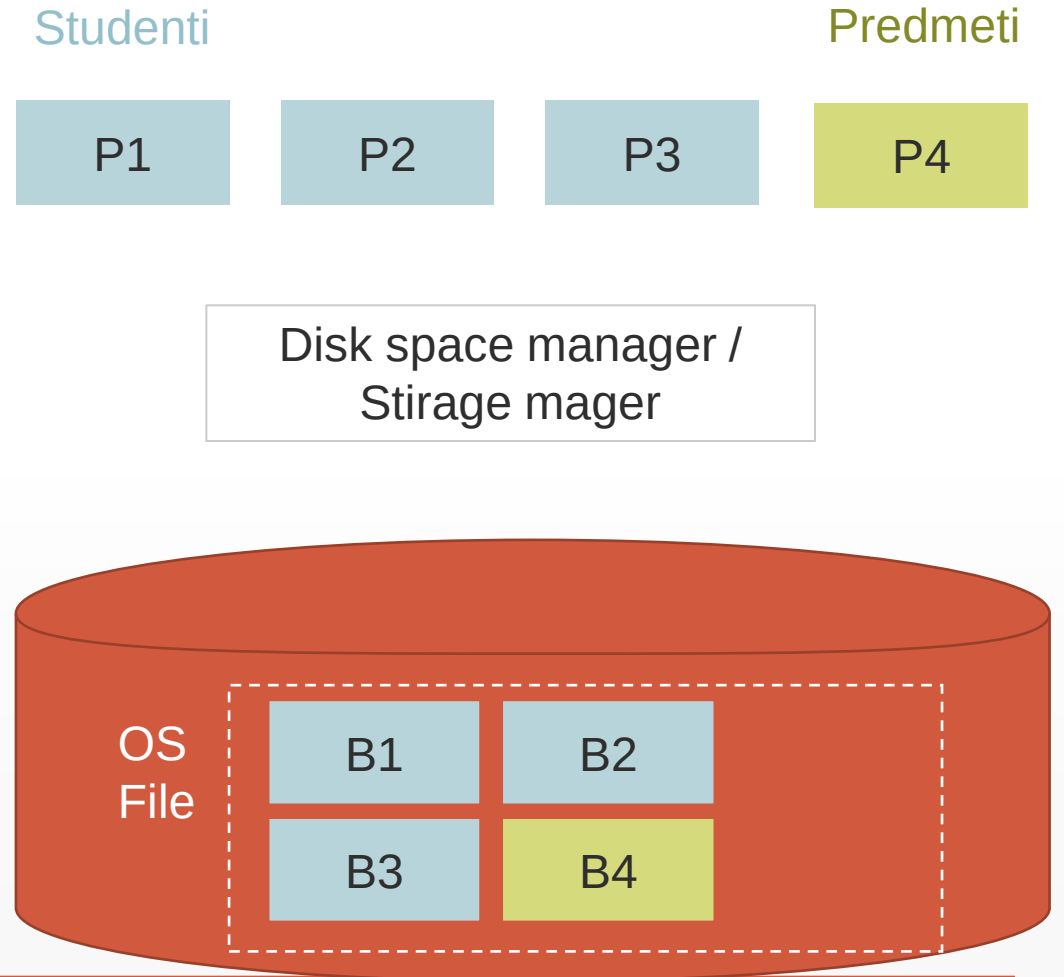


DBMS 'virtuelni' fajlovi

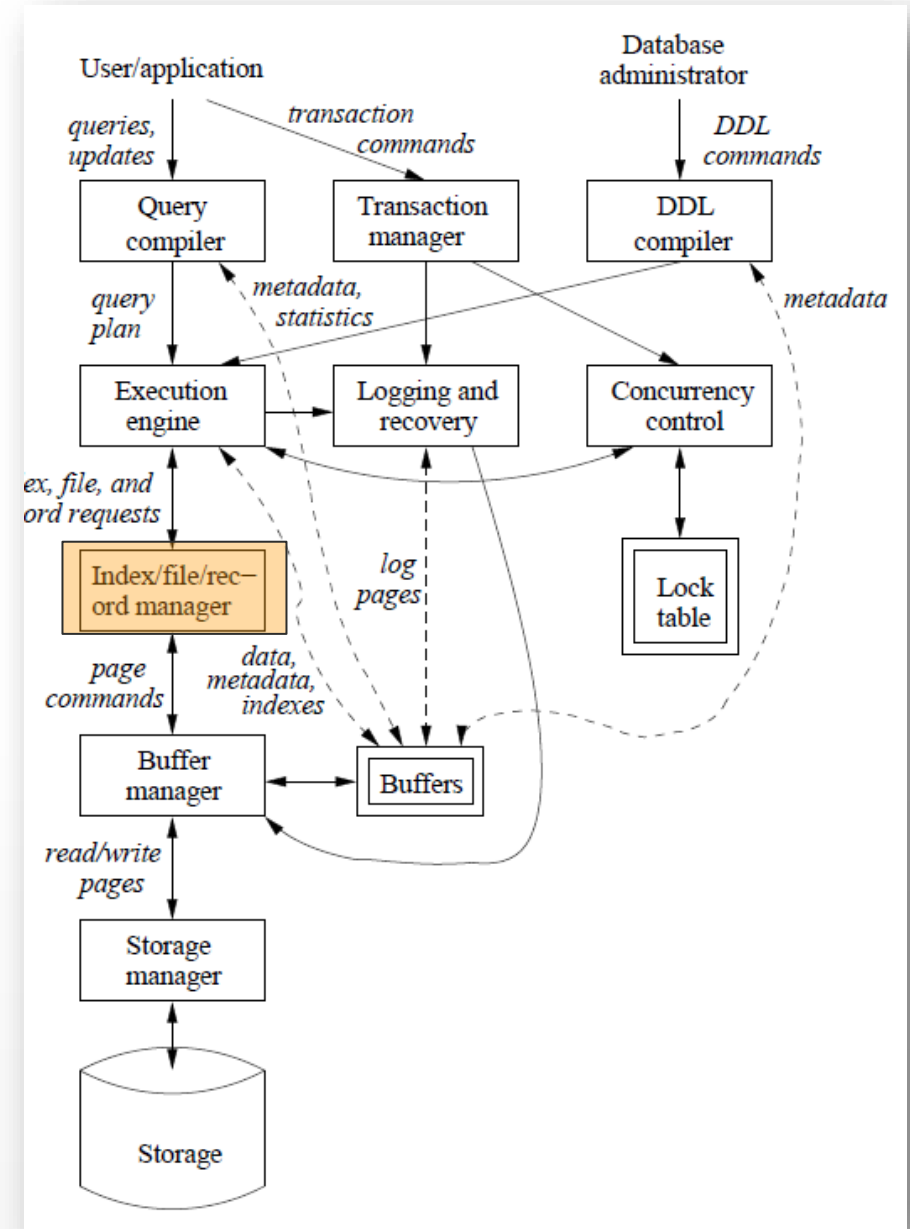
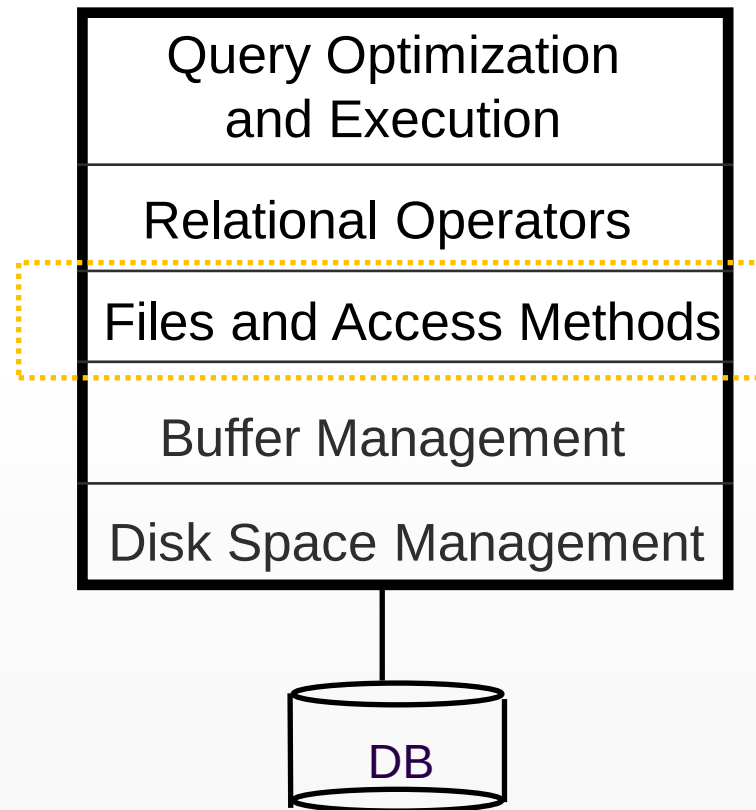
File of records

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

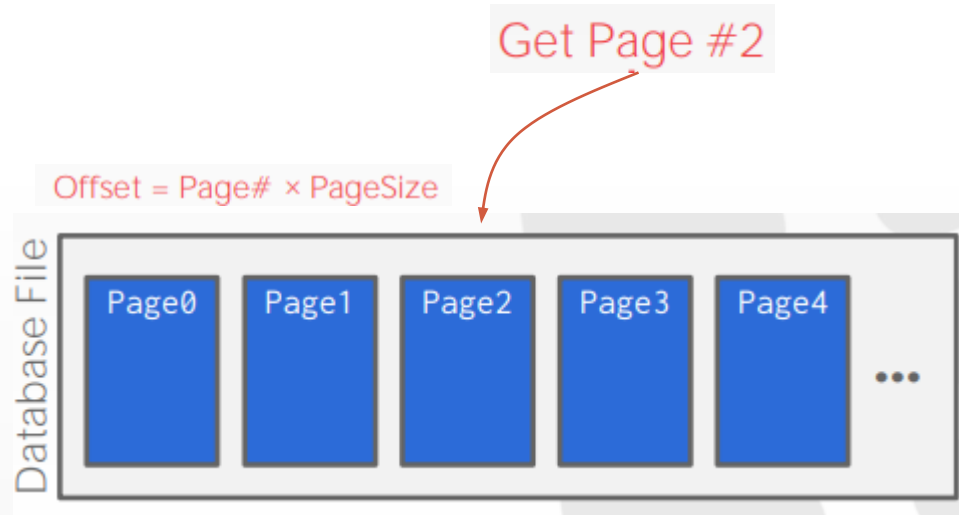
- Menadžer fajlova (index/file/record manager) formira i održava fajl kao kolekciju strana, u kojima se nalaze slogovi.
- Svaka strana
 - je kolekcija slotova, od kojih svaki sadrži po jedan slog
 - svaka strana u fajlu je iste veličine
- Svaki slog (n-torka, record) se identifikuje sledećim parom
(pageID,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

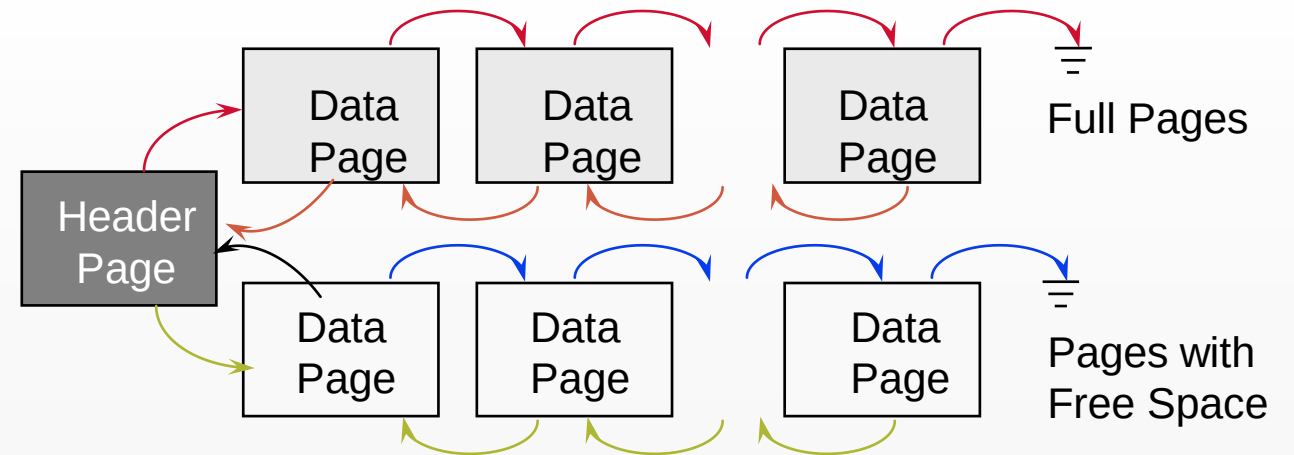
Kako voditi evidenciju stranama koje pripadaju DB fajlu



Heap fajl implementiran kao lista

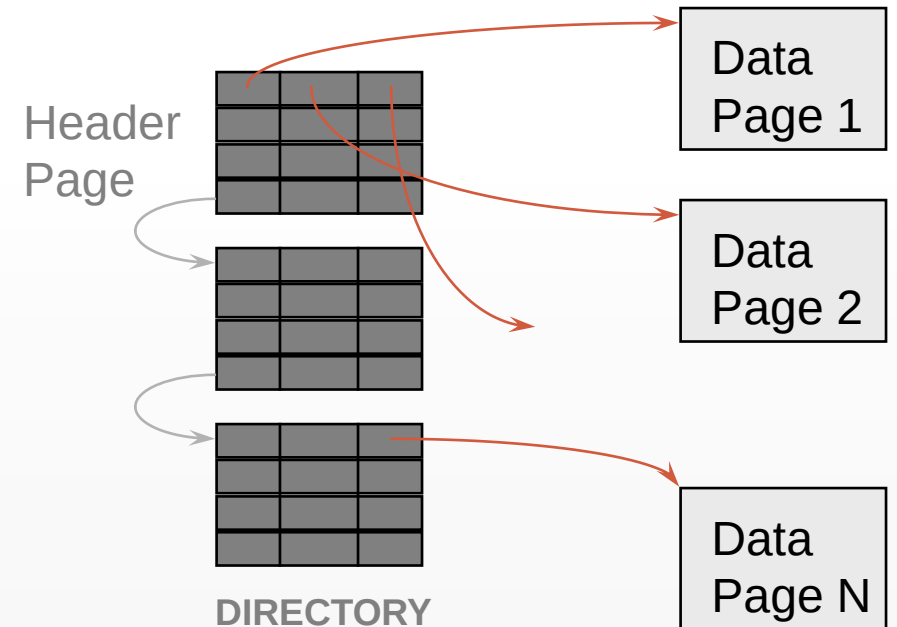
- Da bi obezbedili održavanje informacija o slobodnom prostoru, heap fajl je moguće implementirati kao **dve dostruko povezane liste**, gde je jedna **lista popunjenih strana**, a druga **lista strana sa slobodnim prostorom**.
- Za sve fajlove DBMS pamti informaciju o prvoj strani fajla, tzv. **header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

(heap_file_name, page_Laddr)



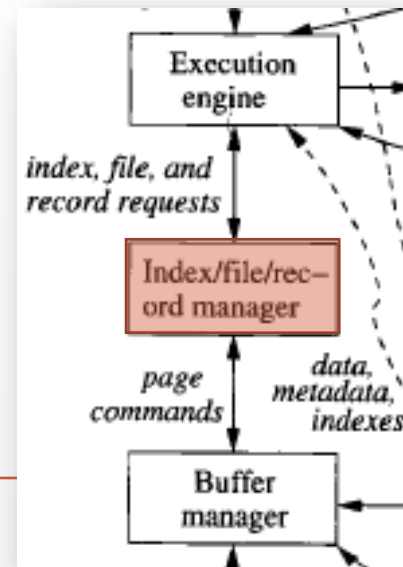
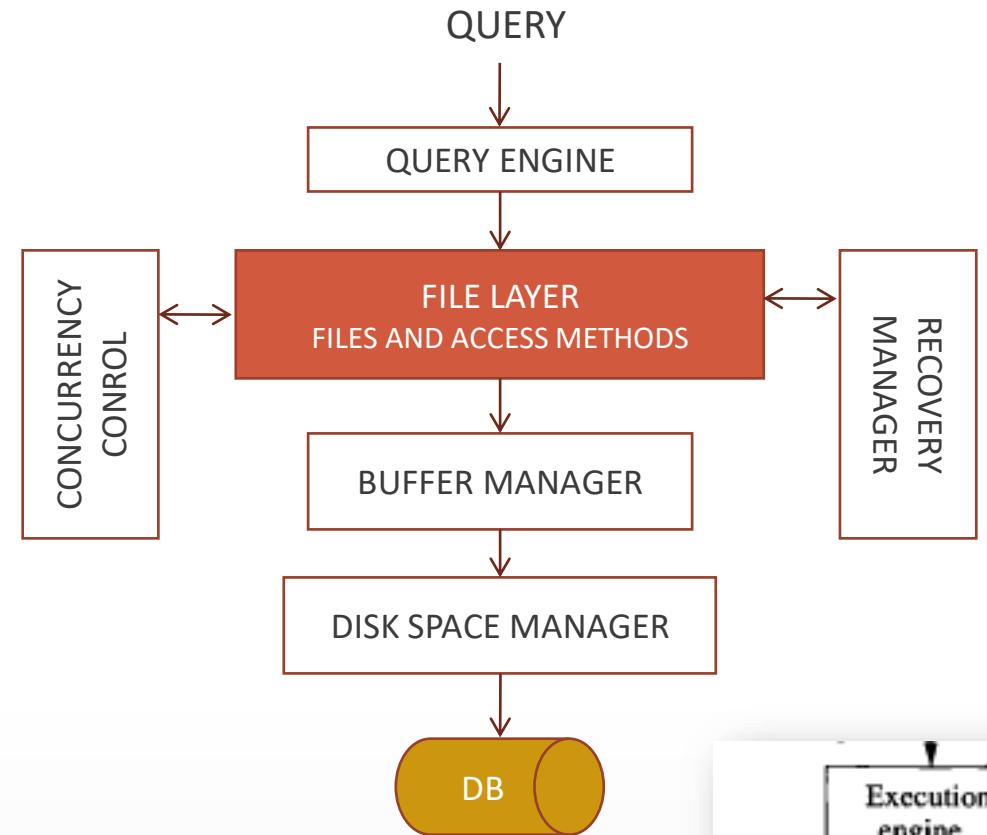
Heap fajl implementiran kao direktorijum strana

- 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



Menadžer fajlova/slogova

- Zahtevi za podacima bafer menadžeru stižu iz **File layer**-a.
- Sloj upravljanja fajlovima:
 - Odlučuje o rasporedu slogova u fajlu,
 - Održava liste strana dodeljenih fajlovima.
 - Vodi evidenciju o slobodnom prostoru na stranama fajla.



Slogovi

Slogovi i polja

- Slog predstavlja kolekciju polja (koja su slogom povezana u logičku celinu)
 - Slog ima:
 - Manje ili više striktno definisanu strukturu / format
 - Polje nosi podatak koji je određenog tipa
 - Pitanja:
 - Kako se beleže pojedini tipovi podataka?
 - Kako se beleže formati?

Dodatno: kako oba utiču na organizaciju slogova u fajlu?
 - Slogovi se beleže kao niz bitova/bajtova. Zadatak DBMSa je da ih protumači i izdvoji pojedine delove kao vrednosti atributa jedne torke.
-

Tipovi podataka / polja – celobrojne vrednosti

- Većina DBMSova beleži celobrojne vrednosti koristeći sopstvene C/C++ tipove prema IEEE-754 standardu
- Vrednosti su fiksne dužine
- INTEGER, BIGINT, SMALLINT, ...

00000000

00100011

Tipovi podataka / polja – realne vrednosti

- Brojevi varijabiljne preciznosti
 - C/C++ tipovi prema IEEE-754 standardu
 - FLOAT, REAL

```
#include <stdio.h>

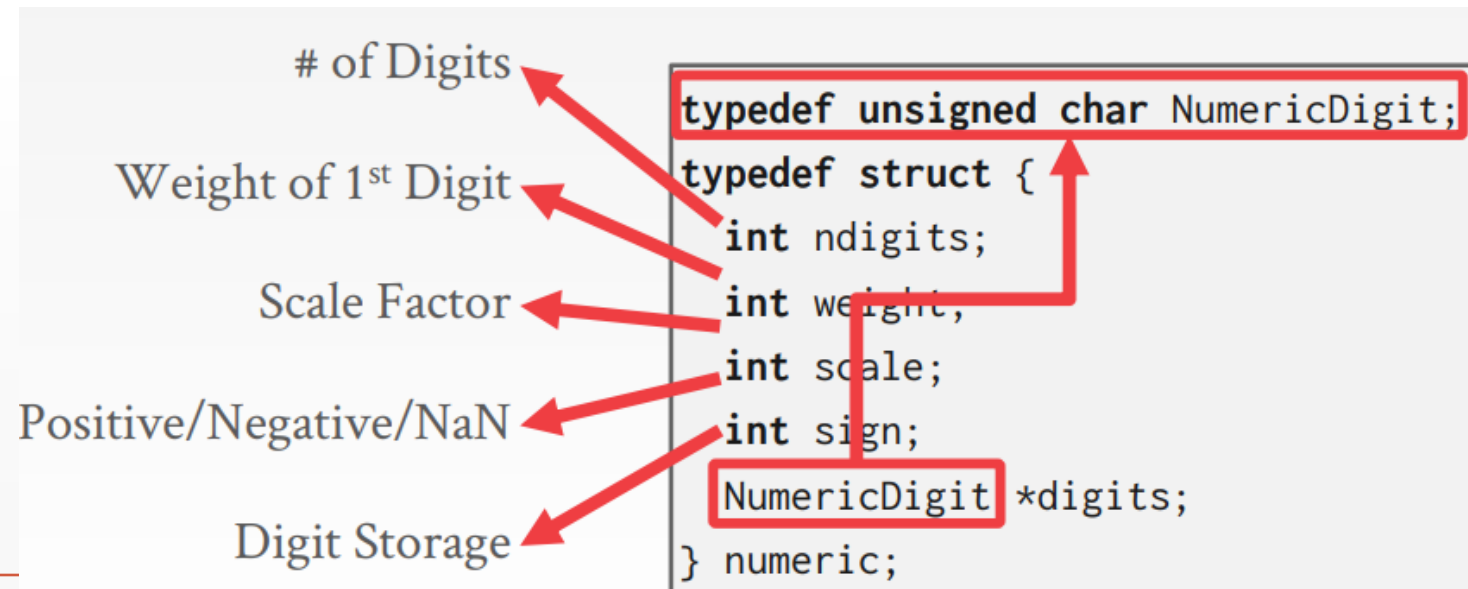
int main(int argc, char* argv[]) {
    float x = 0.1;
    float y = 0.2;
    printf("x+y = %.20f\n", x+y);
    printf("0.3 = %.20f\n", 0.3);
}
```

```
x+y = 0.30000001192092895508
0.3 = 0.299999999999999998890
```

Tipovi podataka / polja – realne vrednosti

- Brojevi fiksne preciznosti
 - Proizvoljna preciznost i opseg
 - Obično sačuvani striktnom reprezentacijom, varijabilne dužine sa dodatnim metapodacima.
 - Operacije nad njima su sporije.
 - NUMERIC, DECIMAL

Postgres



Tipovi podataka / polja – karakteri, bool, stringovi

- Karakteri
 - Koriste se različite kodne šeme
 - Primer ASCII

Example:

A: 1000001

a: 1100001

5: 0110101

LF: 0001010

- Boolean

e.g., TRUE

FALSE

1111 1111
0000 0000

- String of characters

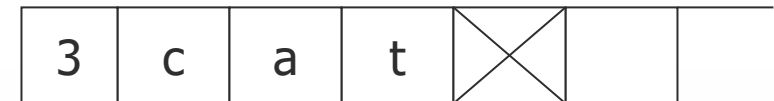
Null terminated

e.g.,



Data dužina

e.g.,



Tipovi podataka / polja – veliki podaci

- Bag of bits

Length	Bits
--------	------

BLOB (Binary Large OBject)

- Većina DBMSova ne dozvoljava slogu da prevaziđe veličinu jedne strane.
 - Oni koji dozvoljavaju koriste overflow stranice (stranice prekoračenja), pri čemu početna svaki deo torke na jednoj strani sadrži referencu na stranicu u koju je smešten nastavak.
 - Neki sistemo dozvoljavaju čuvanje u eksternim fajlovima (na primer slike). Mana ovakvog skladištenja što su izmene nad ovako sačuvanim podacima van kontrole DBMSa (oporavak od otkaza, na primer, ih neće obuhvatiti).
-

Tipovi podataka / polja – datum, vreme

- TIME, DATE; TIMESTAMP
 - Datum
 - e.g.: - Integer, # days since Jan 1, 1900
 - 8 characters, YYYYMMDD
 - 7 characters, YYYYDDD
 - Vreme
 - e.g. - Integer, seconds since midnight
 - characters, HHMMSSFF
-

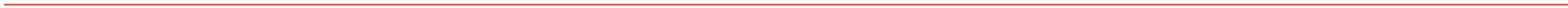
Sistemiški katalog

- Metapodaci o bazama podataka.
 - Čuva podatke o:
 - Tabele, kolone, indeksi, pogledi
 - Kosrisnici, dozvole,
 - Statistike
 - DMBS ima sopstvene kodove za beleženje i čitanje.
-

Tipovi podataka / polja

Sveukupno

- Polje fiksne dužine
- Polje promenljive dužine
češći izbor – dužina data na početku zapisa



Tipovi slogova – Fiksni format

- 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
 - Začenje svakog polja

Employee record

(1) E#, 2 byte integer

(2) E.name, 10 char.

(3) Dept, 2 byte code

} Schema

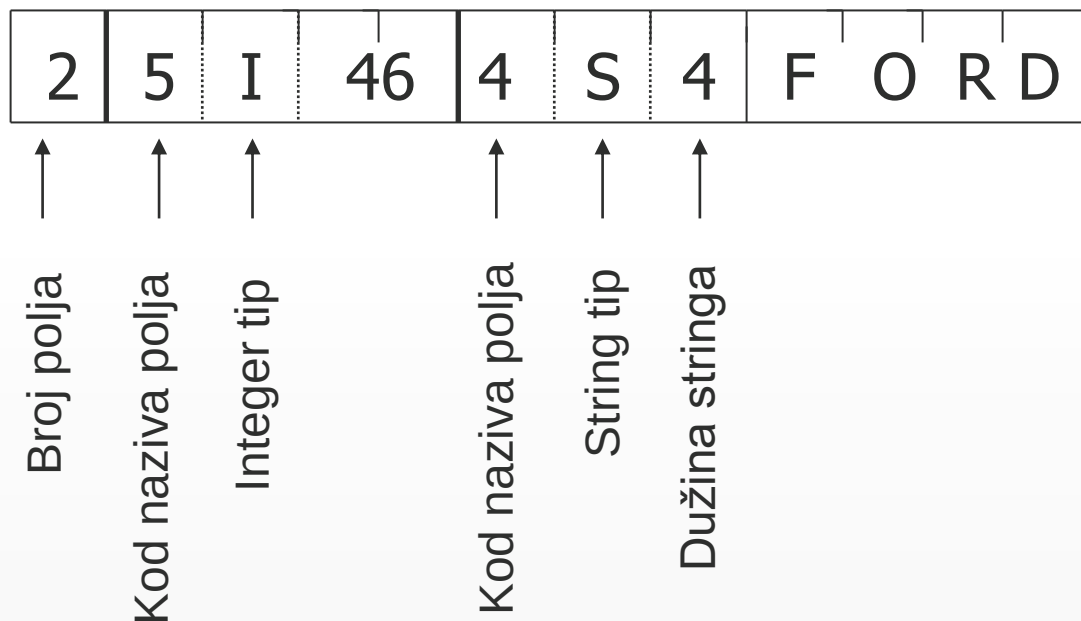
55	s	m	i	t	h					02
----	---	---	---	---	---	--	--	--	--	----

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

} Records

Tipovi slogova – Promenljivi format i dužina

- Samoopisujući format



Kodovi naziva polja mogu biti bilo kakvi tagovi.

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
---	--------------	--------------	------------

Uključivanje tipa sloga u slog

Example: Include record type in record



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

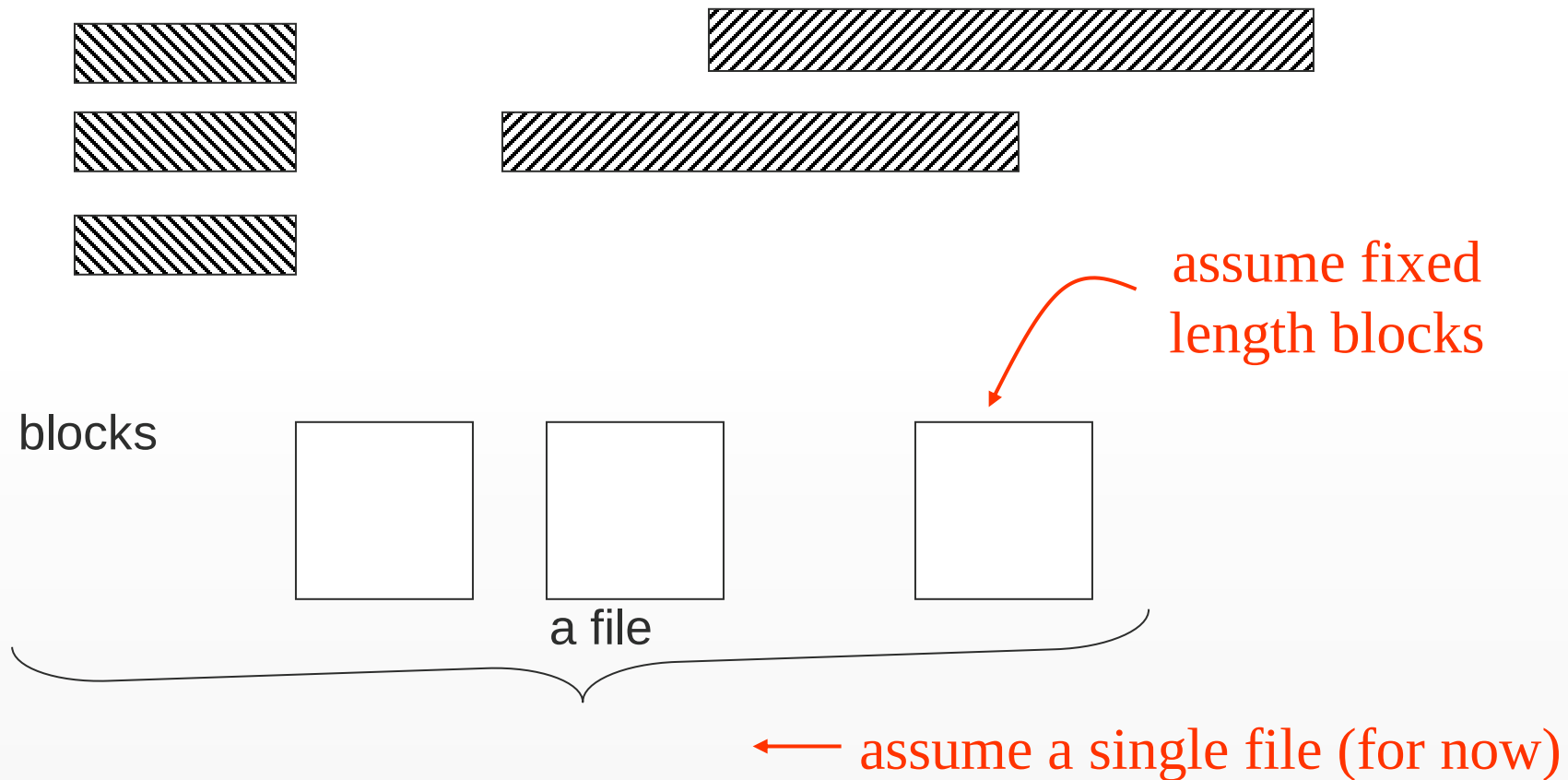
←
record length

Zaglavlje sloga

- Slogovi mogu sadržati zaglavlje u kom se mogu naći
 - Tip sloga (pointer na šemu)
 - Dužina sloga
 - Vremenski pečat
 - ...
-

Beleženje slogova unutar strane

Menadžer fajlova/slogova



Četiri osnovna načina beleženja sloga

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

Razdvajanje slogova



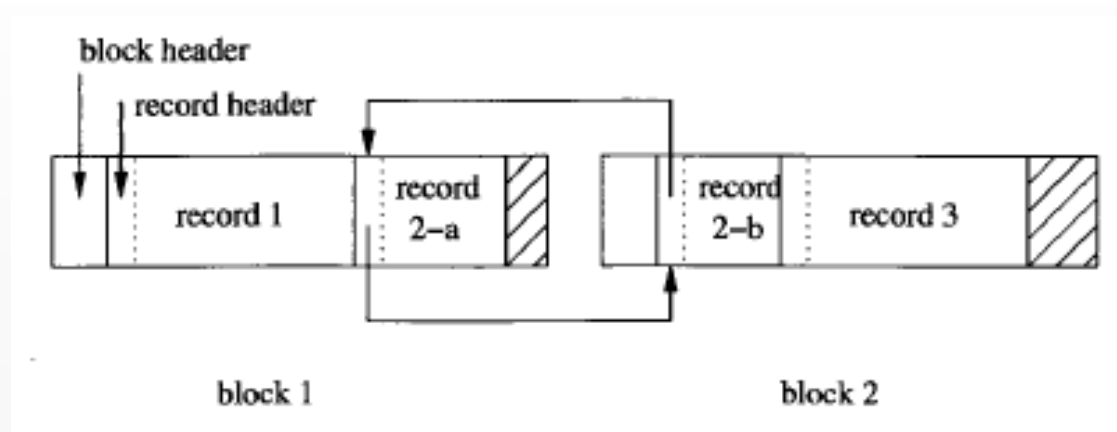
- Slogovi fiksne dužine – ne moraju imati separator
- Uvođenje specijalnog markera kao separatora
- Zadavanje dužine sloga (ili ofseta)
 - Unutar svakog sloga
 - U zaglavlju bloka

Razapeti i nerazapeti slogovi

- Nerazapeti slogovi – jednostavniji, postojanje neiskorišćenog prostora

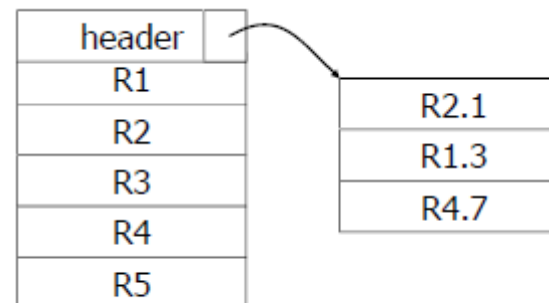
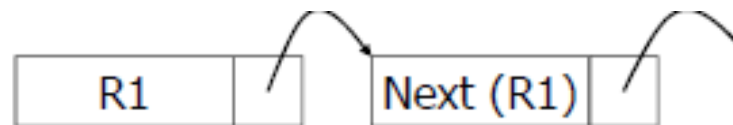


- Razapeti – neophodni kada je veličina sloga veća od veličine sloga



Preusmeravanje

- 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

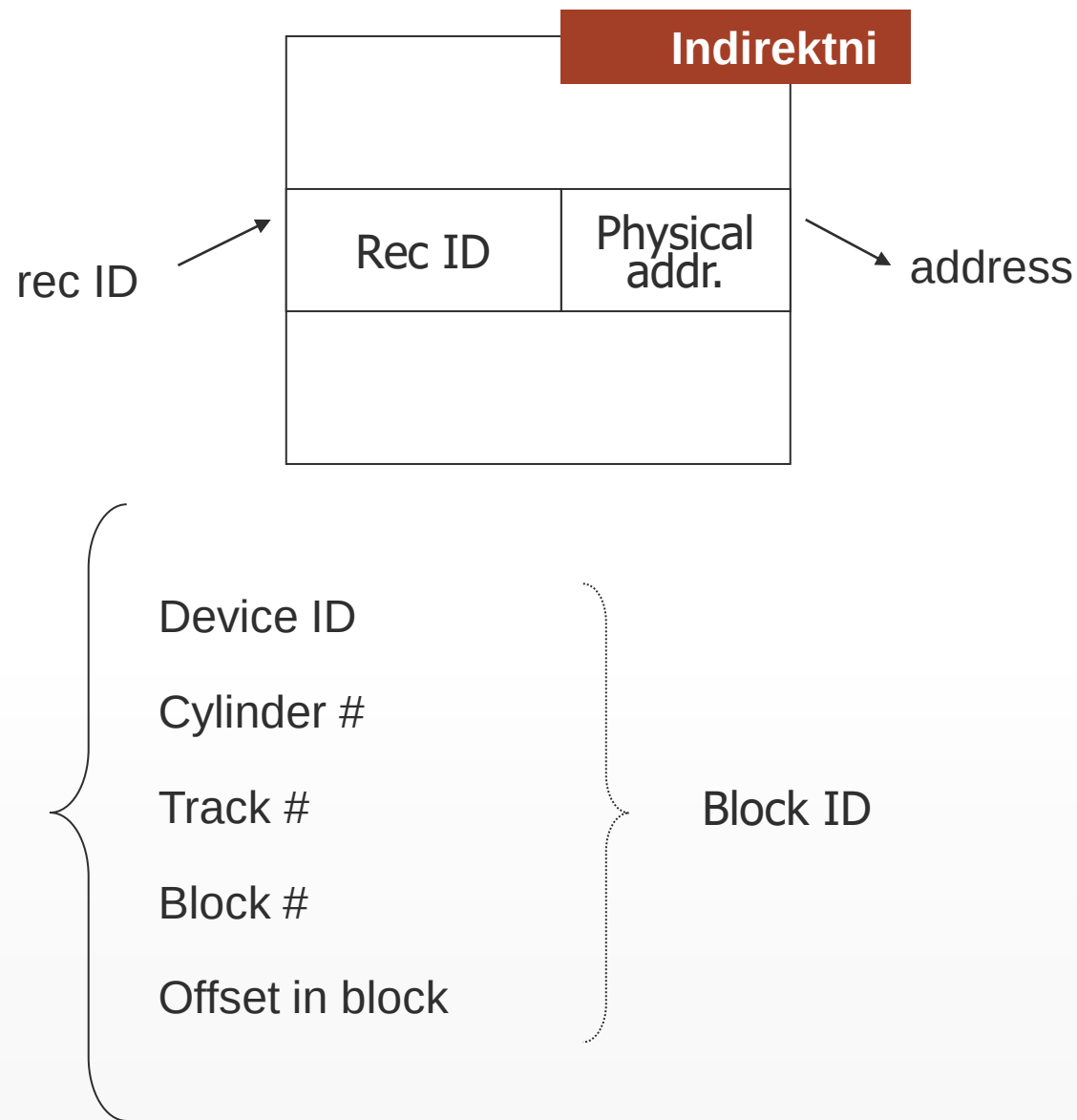


Preusmeravanje

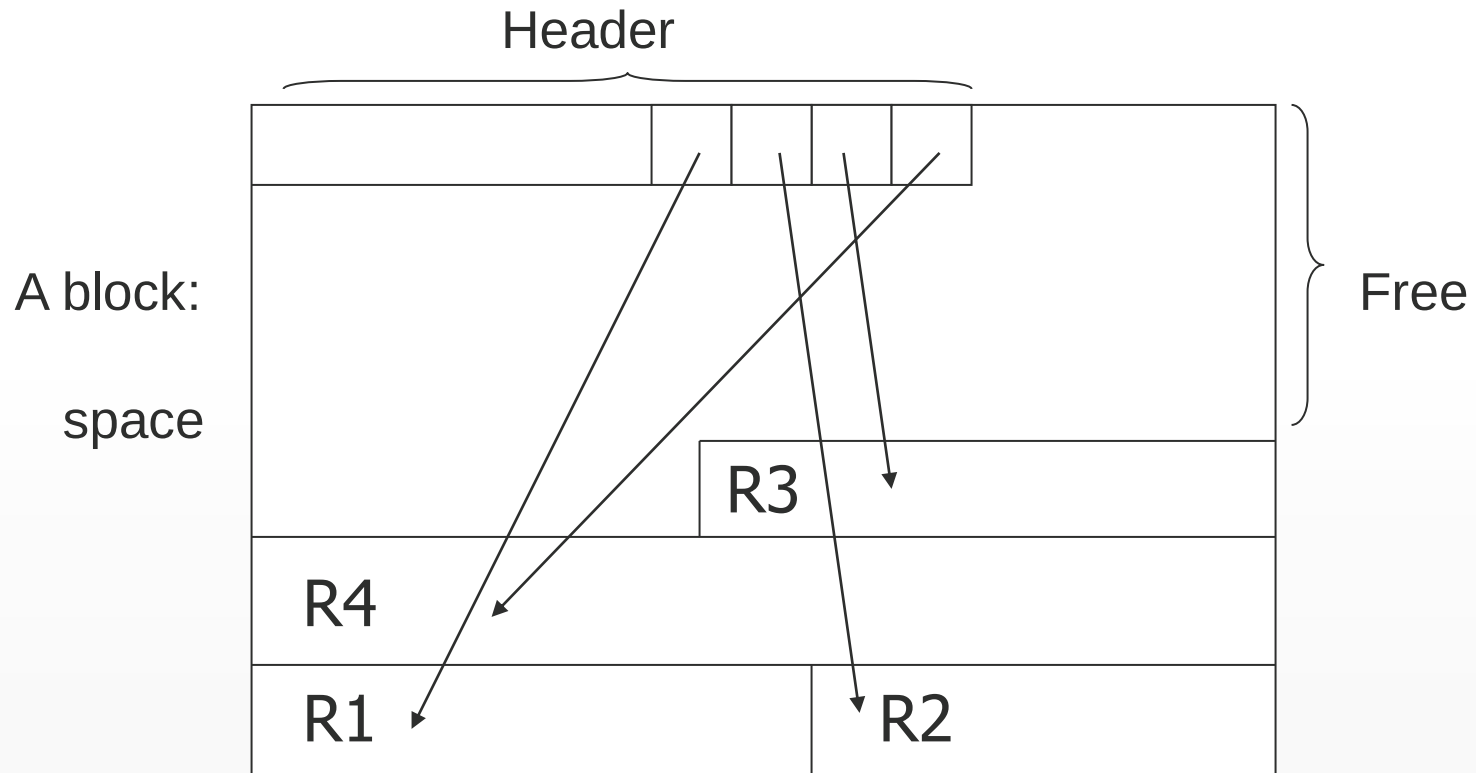
- Pokazivač na slog može biti:
 - Fizički
 - Indirektni

E.g., Record
Address =
or ID

Fizički



Primer preusmeravanje



Organizacija unutar strane

Logička organizacija

Strana kao kolekcija slogova

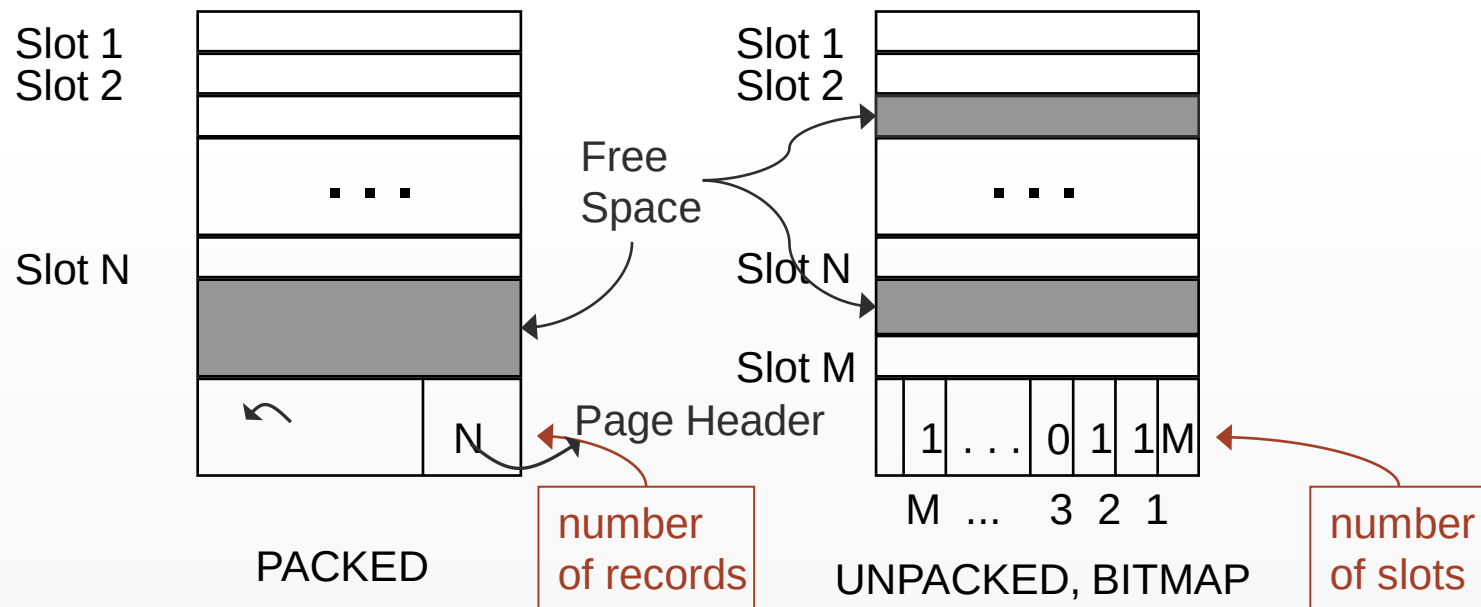
- Strana se može posmatrati kao kolekcija slotova, gde svaki slot sadrži jedan slog.
 - Svaki slog je idetifikovan parom **(ID strane, broj slota) – rid**.
 - Organizacija slogova unutar strane se ocenjuje sa aspekta sprovođenja operacija pretrage, dodavanja i brisanja slogova.
-

Zaglavlje strane

Može da sadrži

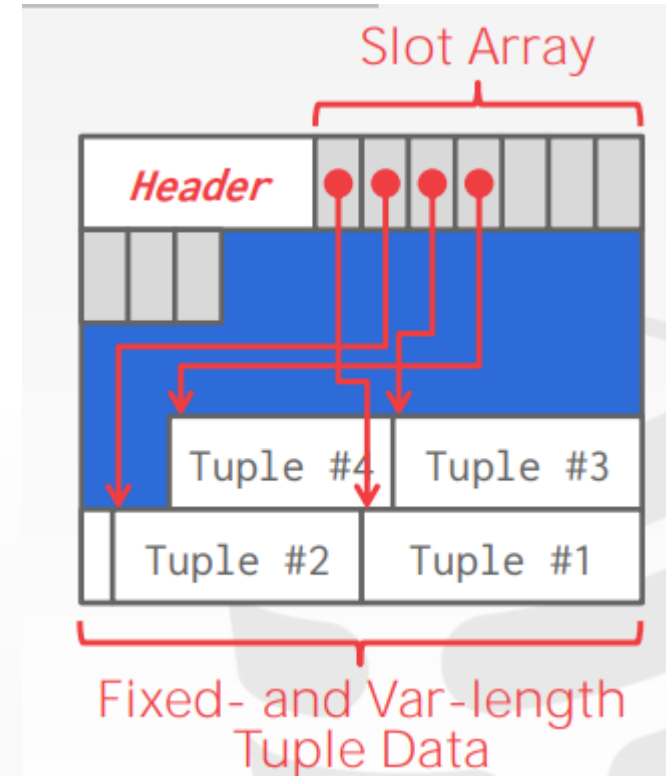
- File ID (RELATION or DB ID)
 - Blok ID
 - Tabela/directory slogova
 - Pointer na slobodan prostor
 - Tip bloka (npr. sadrži slogove tipa 4, blok prekoračenja)
 - Vremeski pečat
-

Strana koja sadrži slogove fiksne dužine



Strana koja može da sadrži slogove promenljive dužine

- Najfleksibilniji način organizacije strane sa slogovima promenljive dužine je održavanje tabele slotova, gde se svaki slot opisuje parom
(record offset, record length)
- Offset – broj bajtova od početka prostora sa podacima.
- Brisanje – postavljanje offseta na -1.
- Slot kojem je pridružen obrisati slog ne može uvek biti obrisani, jer bi poremetio brojanje, tj. redne brojeve slotova, a time i redove postojećih slogova.
- Može da čuva slogove fiksne i varijabiljne dužine.



Strana koja može sadrži slogove promenljive dužine

