

# **Введение в реляционные базы данных**

Лекция 4: Хранение данных, часть 2

Артем Толканев

October 16, 2024

# На этой лекции

- Продолжим изучать хранение данных



## Хранение СУБД

На диске присутствует файл базы данных (или несколько)



Файл  
БД



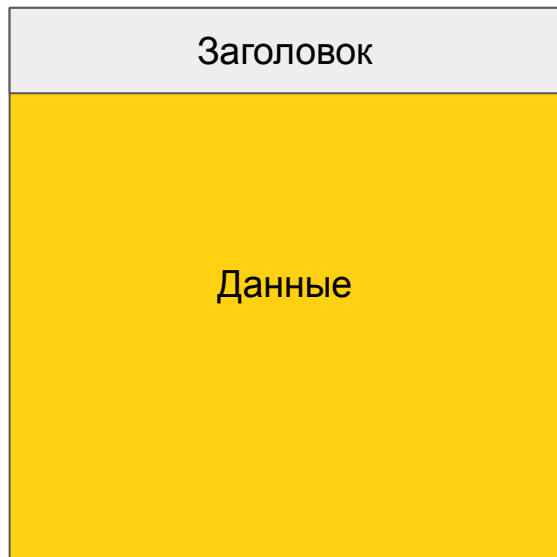
Файл базы данных состоит из нескольких частей, с которыми работает СУБД



# Страницы

Заголовок
Данные





У разных СУБД разные данные в заголовке и разная длина заголовка

1. Checksum
2. Ссылка на последнюю измененный кортеж в данных
3. Количество свободного места на странице



Заголовок			
1	2	3	Данные
			Кор
теж 3	Кортеж 2	Кортеж 1	

Массив с информацией о расположении кортежей

Заголовок		
Кортеж 1	Кортеж 2	Кор
теж 3	Данные	
		3
		2
		1





Postgres

Массив с  
информацией о  
расположении  
кортежей



MS SQL, MySQL



Массив с  
информацией о  
расположении  
кортежей

```
SELECT  
sys.fn_PhysLocFormatter(%%physloc%%),  
book_id  
FROM book  
  
(file_id:page_id:row_id)
```



MS SQL







Массив с  
информацией о  
расположении  
кортежей

```
SELECT ctid, book_id  
FROM  
book
```

Postgres



# Строки

Заголовок		
Кортеж 1	Кортеж 2	Кор
теж 3	Данные	
		3 2 1

Заголовок	Данные
-----------	--------

1. Информация о NULL
2. Информация транзакционная
3. Информация о данных фиксированной\переменной длины



# Пауза на DDL

**SELECT** \* **FROM** book

book_id	name	theme_id	publishing_house_id	year_of_publishing	language_id
200	Война и Мир 1,2 том	1	11	1 998	1
201	Война и Мир 3,4 том	1	11	1 998	1
202	Othello	2	12	2 005	2
203	Курс аналитической геометрии	5	13	2 005	1
204	Обломов	1	14	2 005	1
205	Капитанская Дочка	1	15	2 005	1
206	Общая физика	5	13	2 005	1
207	Дубровский	1	15	2 005	1
208	Анна Каренина	1	11	2 005	1
209	1984	2	12	2 005	2
210	Мартин Иден	2	16	2 005	1
211	Сердца Трех	2	16	2 007	1
212	Белый Клык	2	16	2 007	1
213	Три Сестры	1	14	2 007	1
214	Русские сказки	4	14	2 007	1
215	Курс аналитической геометрии	5	13	2 006	1



# Пауза на DDL

```
CREATE TABLE book (  
    book_id int PRIMARY KEY,  
    name varchar(50) NOT NULL,  
    theme_id int NOT NULL,  
    publishing_house_id int NOT NULL,  
    year_of_publishing int NOT NULL,  
    language_id int NOT NULL  
);
```



```
CREATE TABLE book (  
  book_id int PRIMARY KEY,  
  name varchar(50) NOT NULL,  
  theme_id int NOT NULL,  
  publishing_house_id int NOT NULL,  
  year_of_publishing int NOT NULL,  
  language_id int NOT NULL  
);
```

Первичный ключ

все значения в строке  
должны быть не NULL



Ограничение **UNIQUE** гарантирует, что все значения в столбце будут разными.

```
CREATE TABLE bucket (  
    id int PRIMARY KEY  
    , name varchar(1) UNIQUE  
);
```

Вопрос: у нас есть способ вставить данные в таблицу `INSERT INTO bucket VALUES ...`. Как поведет СУБД при следующей команде:

```
INSERT INTO bucket VALUES  
(1, NULL)  
, (2, NULL)
```



# Строки

Заголовок		
Кортеж 1	Кортеж 2	Кор
теж 3	Данные	
	3	2
		1

Заголовок	Данные
-----------	--------



Заголовок			
Кортеж 1	Кортеж 2	Кор	
теж 3	Данные		
	3	2	1

Заголовок	Данные
-----------	--------

book_id	name	theme_id	publishing_house_id
---------	------	----------	---------------------





Заголовок			
Кортеж 1	Кортеж 2	Кор	
теж 3	Данные		
	3	2	1

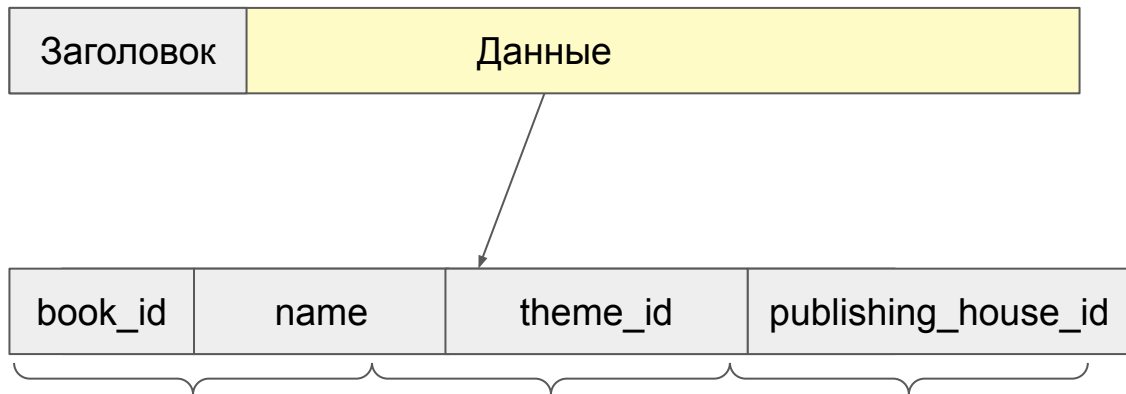
Заголовок	Данные
-----------	--------

book_id	name	theme_id	publishing_house_id
---------	------	----------	---------------------

а здесь что может не так пойти



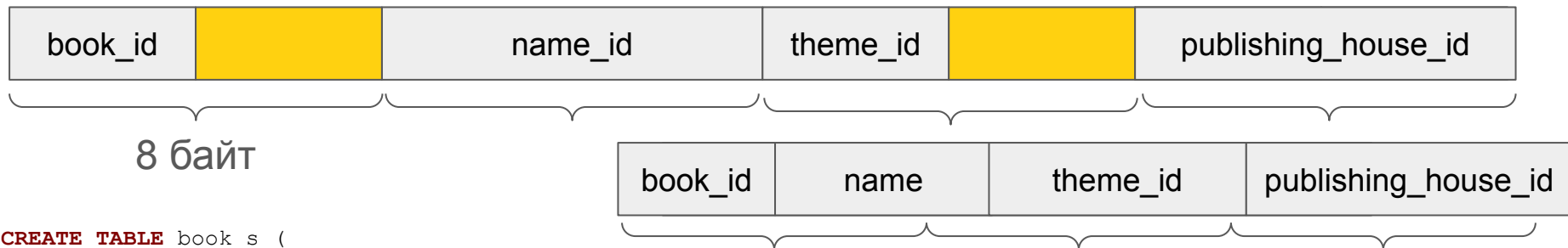
Заголовок		
Кортеж 1	Кортеж 2	Кортеж 3
Данные		
	3	2
		1



хотелось бы чтобы процессор получал доступ к значению без дополнительных работ



Добавьте пустые биты после атрибутов, чтобы убедиться, что кортеж выровнен по словам.



```
CREATE TABLE book_s (  
  book_id int4 NOT NULL,  
  name_id int8 NOT NULL,  
  theme_id int4 NOT NULL,  
  publishing_house_id int8 NOT NULL  
);
```

хотелось бы чтобы процессор получал доступ к значению без дополнительных работ



Типы данных:

int, smallint, bigint и тд.

float/real/double - IEEE 754 стандарт

```
SELECT cast(0.1 AS real) + cast(0.2 AS real)
```



## Типы данных:

int, smallint, bigint и тд.

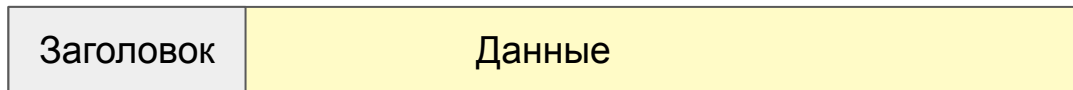
float/real - IEEE 754 стандарт

numeric/decimal - числа с фиксированной запятой

time/date/timestamp - секунды с 1970 года 1 января



# NULLs



1. Информация о NULL в заголовке строки
2. Специальное значение у типа данных, например INT32\_MIN
3. Ставить флаг перед каждым атрибутом

