



Blurred Education

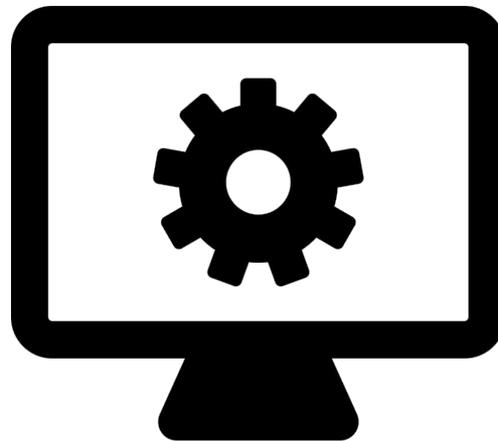
Хранение данных - основы

Булатов Александр

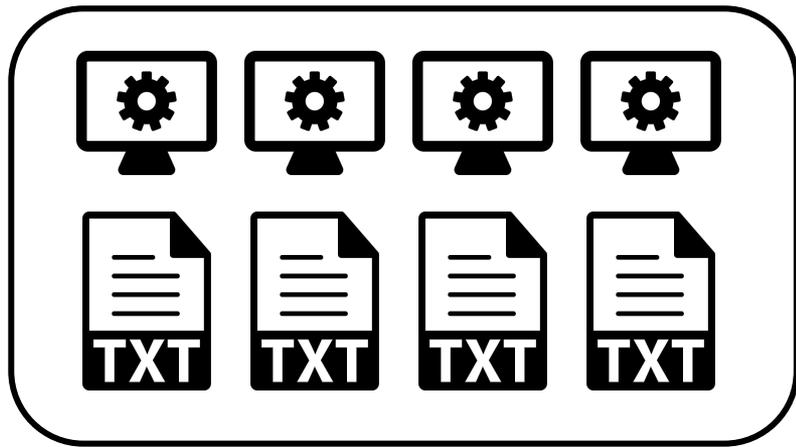
Определяй правильно!



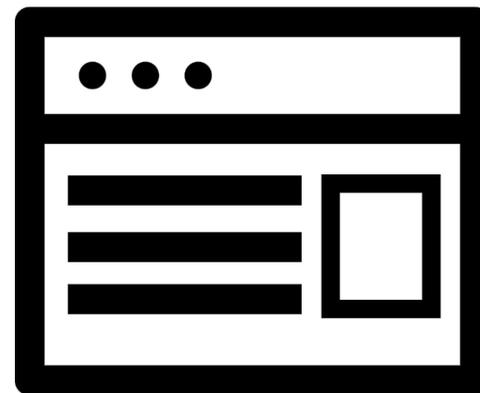
БД
База данных



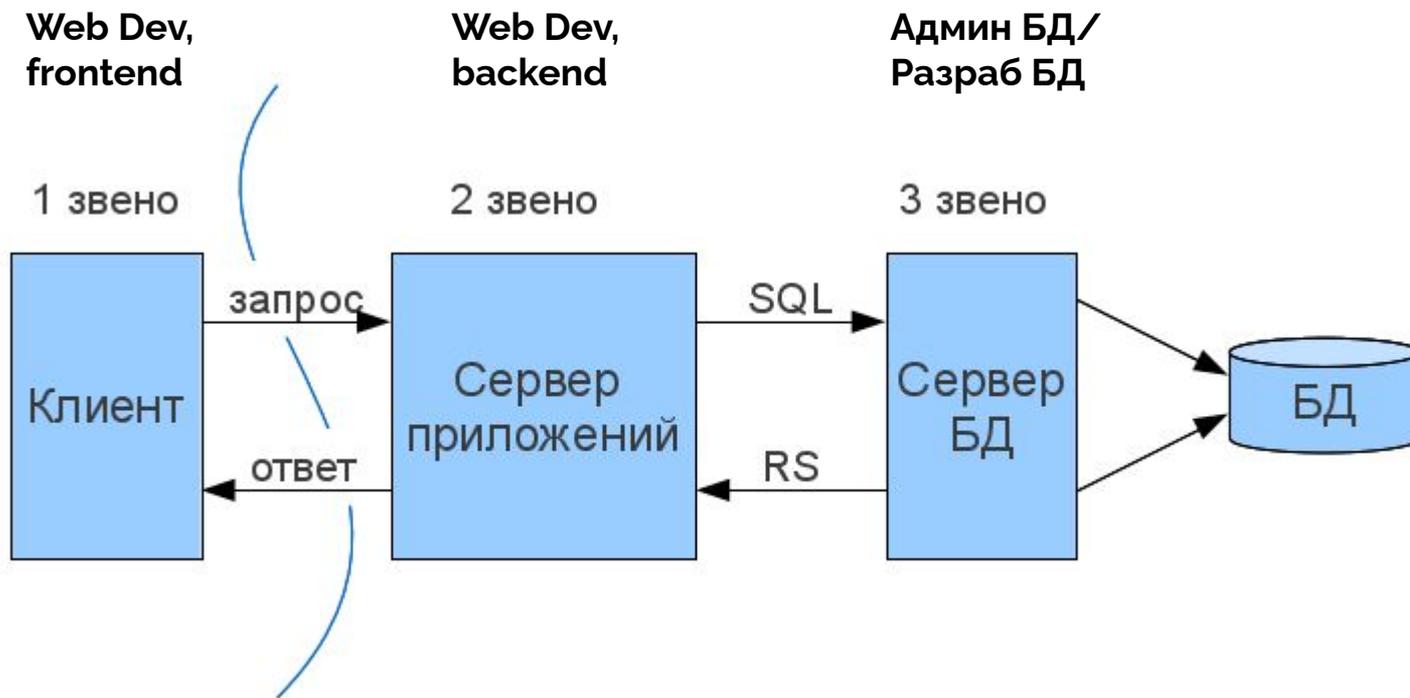
СУБД
Система управления базами данных



Хранилище данных



Витрина данных



- **Администратор баз данных**

Знание работы сетей, SQL, конкретная СУБД, Linux-администрирование

- **Разработчик баз данных**

SQL, ЯП, алгоритмы и структуры данных, принципы проектирования БД, аналитический склад ума, PL/SQL

- **Инженер-программист**

ЯП, аналитический склад ума, SQL, PL/SQL, ORM

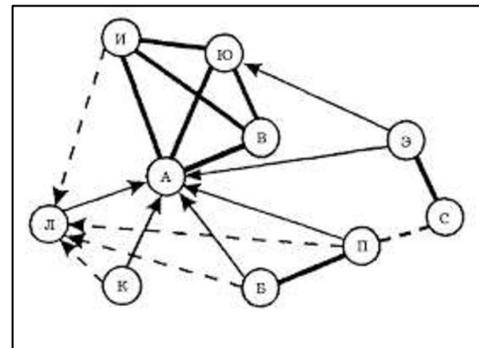
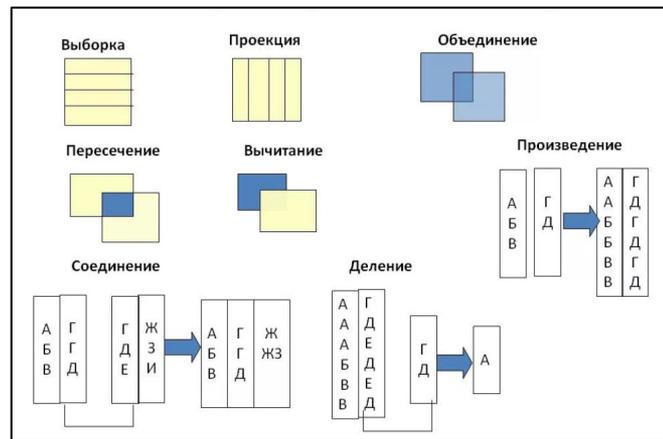
- **Аналитик данных**

SQL, аналитический склад ума, знания в оптимизации запросов, UML, принципы проектирования БД

- **Инженер данных**

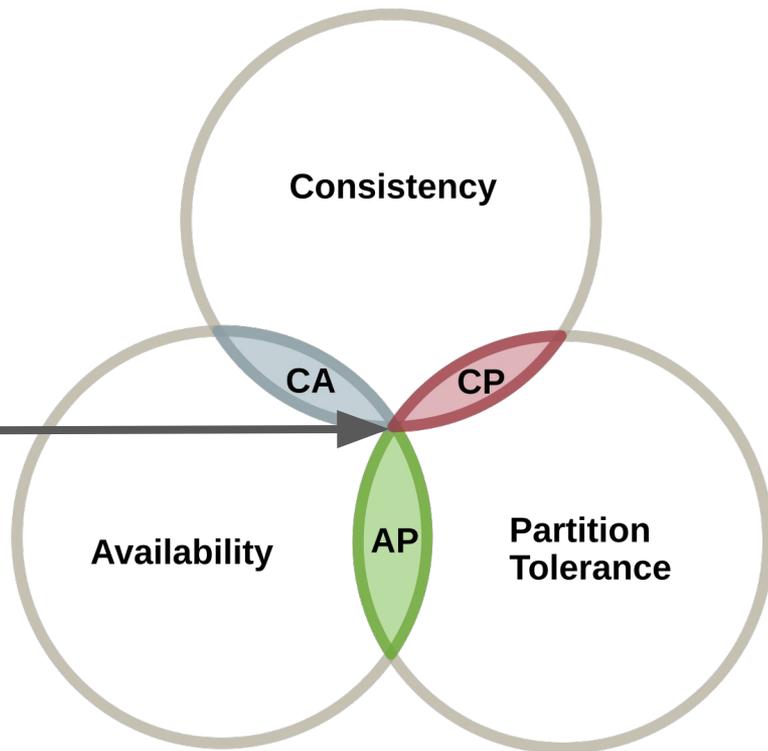
SQL, NoSQL, ЯП (преимущественно Java, Python),

- Дискретная математика
 - Булева алгебра
 - Теория графов
 - Теория множеств
 - Реляционная алгебра
 - Основные операции
 - Нормальные формы
- Аналитическое и инженерное мышление
- (Позже) Навыки Python и других ЯП



Теорема CAP

- Consistency - согласованность
- Availability - доступность
- Partition tolerance - устойчивость



- PostgreSQL - основная СУБД на направлении
- MongoDB - NoSQL база данных
- Python/C/C++/C#/Java - языки разработки ПО для связи "интерфейса с данными"
- DBeaver, DataGrip - инструменты для разработки и проектирования БД
- ClickHouse - NoSQL колоночная база данных от Яндекса (эксперименты)
- Tarantool - экспериментальная СУБД на направлении, NewSQL от Mail.ru

SQL - Структурированный язык запросов, основан на **реляционной** модели данных.

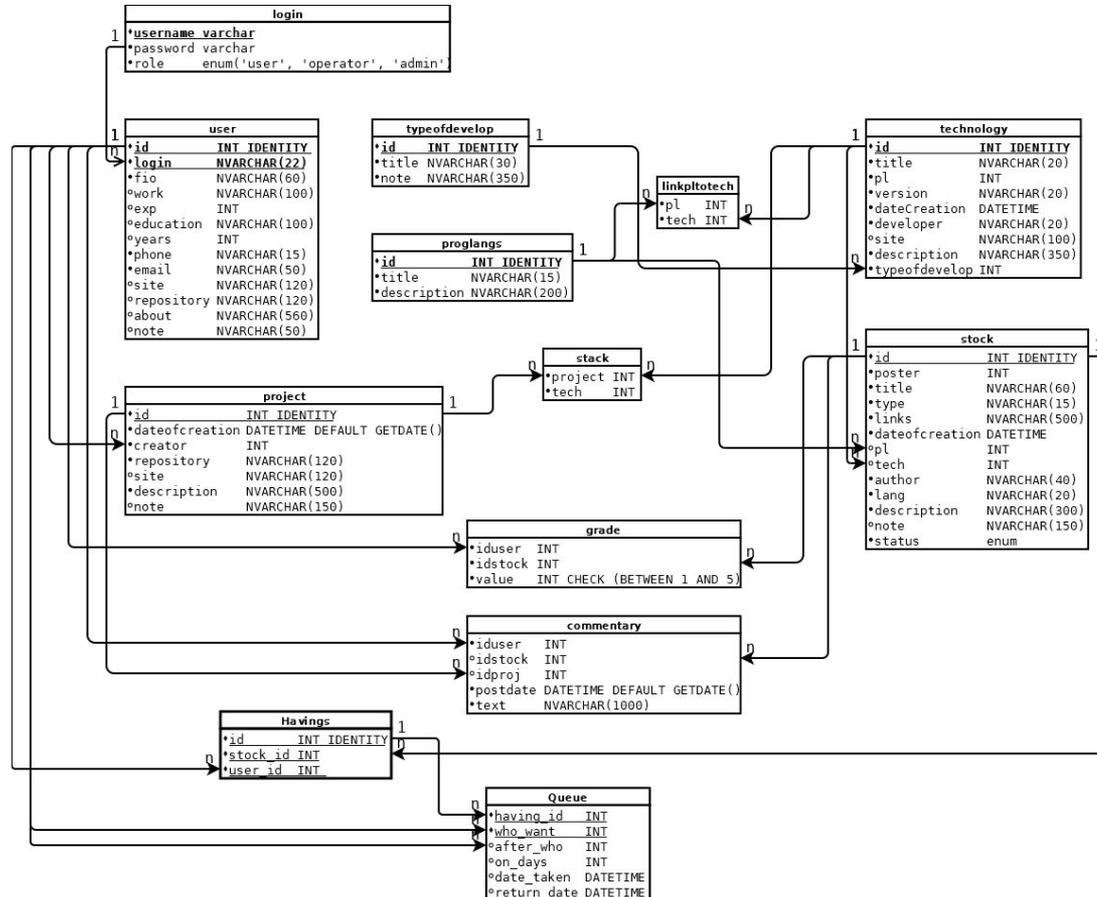
```
CREATE TABLE `table1` (`user_id` INT(5) NOT NULL AUTO_INCREMENT,  
`username` VARCHAR(50), PRIMARY KEY(`user_id`), INDEX(`username`));
```

```
SELECT phone_id, user_id FROM table2 WHERE phone_number=200
```

```
SELECT phone_id, user_id FROM table2 WHERE phone_number=200 LIMIT 1, 3;
```

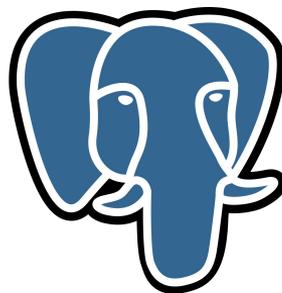
```
DELETE FROM table1 WHERE user_id = 1;
```

Больше - в постах и воркшопе



SQL имеет помимо стандартов разные диалекты, которые определяют разработчики СУБД.

- Oracle SQL
- MySQL
- PostgreSQL
- Microsoft SQL Server
- SQLite



Целевые системы

- Корпоративные системы
- Банковское и финансовое ПО
- Научные базы данных
- Крупные системы
- Телекоммуникационные системы

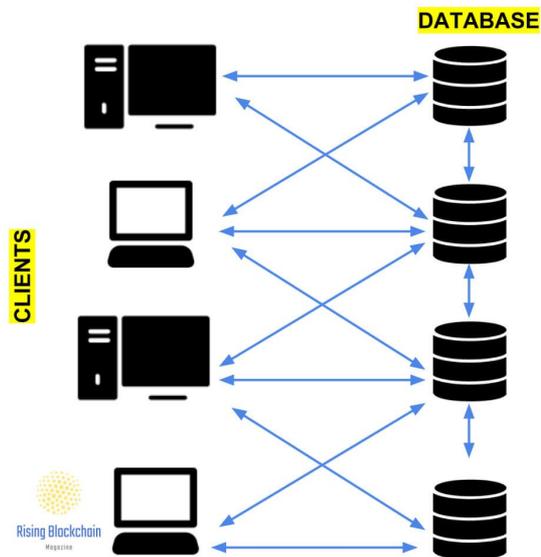
Задачи и инструменты:

- SQL
- PL/SQL
- Разработка хранилища данных
- Администрирование баз данных
- Оптимизация запросов
- Разработка распределенных систем
- Анализ данных

Проблемы SQL СУБД:

- Сильноструктурированные данные
- Медленные запросы
- Мало возможностей для распределенной архитектуры

Решение: **Not Only SQL (NoSQL)**



NoSQL - СУБД, не придерживающиеся принципов SQL систем (ACID)

Опирается на масштабируемость за счет атомарности и
консистентности данных

Виды:

- Key-Value - хранилище - Redis
- Колоночные БД - Google BigTable, Clickhouse, Cassandra
- Документоориентированные БД - CouchDB, MongoDB
- Графовые БД - Neo4j



Новый виток развития СУБД

(Wiki) **NewSQL** - класс современных **реляционных СУБД**, стремящихся совместить в себе преимущества **NoSQL** и **транзакционные требования** классических баз данных (ACID)

1. VoltDB - масштабируемая реляционная СУБД
2. PostgreSQL 10 и выше
3. Плагины и расширения для существующих СУБД

1. SQL для простых смертных
2. Практика проектирования БД и запросов к данным
3. СУБД и программы



А ТЕПЕРЬ

ПОГОВОРИМ?