**Базы данных и SQL**

База данных (БД) – структурированное хранилище информации, организованные таким образом, чтобы обеспечить хранение, управление и доступ к информации. Чаще всего эти данные взаимосвязаны. Сама по себе БД – это просто хранилище данных, остальные свойства базы данных обеспечиваются за счет системы управления базой данных.

Система управления базой данных (СУБД) – программные средства для управления данными. Отвечает за поддержку языка БД (обеспечивает интерфейс для взаимодействия с данными), за механизм хранения и извлечения данных, за оптимизацию процессов извлечения данных.

**Виды БД:**

*Без использования СУБД*

|  |  |
| --- | --- |
| Файловые | Данные хранятся в виде текста в файле. |
| Иерархические | Имеет древовидную структуру (узел может иметь только одного родителя) Пример: файловая система. |
| Сетевые | Имеет структуру графа (узел может иметь более одного родителя) |

*С использованием СУБД*

|  |  |
| --- | --- |
| Реляционные | Структура в виде таблиц, связанных между собой |
| NoSQL | Общее название для нереляционных БД. |

**Типы СУБД:**

NoSQL (not only SQL) – Нереляционные СУБД, в которых данные могут быть организованы в виде иных структур, отличных от таблиц, ориентированных под свои задачи.

Используется язык отличный от SQL. Отсутствие жесткой схемы данных. Отказ об атомарности и согласованности данных в пользу горизонтального масштабирования.

1) Ключ-значение – база данных организована в виде словаря. Некая таблица, имеющая уникальный ключ и соответствующее ему значение. Подходит для кэширования. [Redis]

Высокая скорость, легкое масштабирование. Не подходит для сложных структур (сущности с разными типами данных) и сложных запросов.

2) Колоночные – Данные организованы в виде таблиц так же, как и в реляционной СУБД, но данные организованы не построчно (для чтения колонки придется вычитать всю строку до нужной колонки), а «поколоночно», т.е. колонка — это отдельная таблица. В отличие от реляционных не имеют жесткой схемы, то есть, хранит только имеющиеся данные (см. рис.2) [HBase, Cassandra]

За счет своей структуры

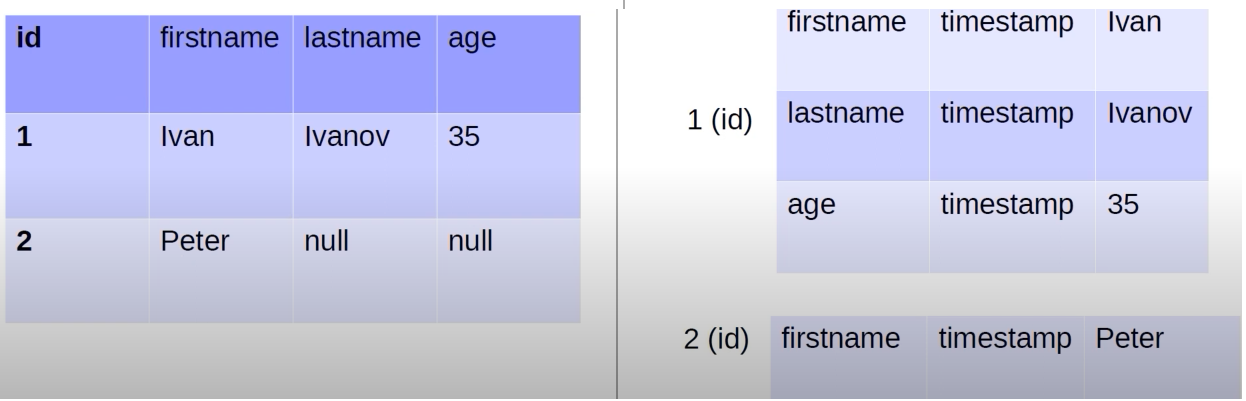
Подходит для аналитических запросов на больших объемах данных (эффективна при запросах для более 100 миллионов строк), позволяет менять схему таблиц – не требует перестройки всей таблицы и не хранить null значения, лучшее сжатие. Cassandra – легко горизонтально масштабируема

Плавающая скорость записи – зависит от архитектурных особенностей и методов записи, но в среднем ниже, чем у реляционных

Общая структура:



Реляционная СУБД Колоночная СУБД



Документо-ориентированные

Базы данных временных рядов

Графовые базы данных

реляционная алгебра

Файл-серверные, клиент-серверные, встраиваемые субд????

Теорема брюйера

ACID