

[Презентация к уроку 6.2.3](#)

Текстовая расшифровка видео:

ЗНАКОМСТВО С NOSQL. ВИДЫ И ПРИМЕНЕНИЕ

План:

- Виды NoSQL БД;
- Колоночные хранилища;
- Графовые базы данных;
- Документоориентированные хранилища;
- Хранилища «ключ-значения»;
- Time-Series DB;
- Виды по назначению;
- Виды по хранению.

Виды NoSQL БД

Виды NoSQL БД делятся:

- По их шаблонам доступа к данным;
- По применению OLTP или OLAP;
- По хранению.



- Колоночные хранилища
- Графовые базы данных
- Документоориентированные хранилища
- Хранилища «ключ-значения»
- Time-Series DB

Шаблоны доступа к данным

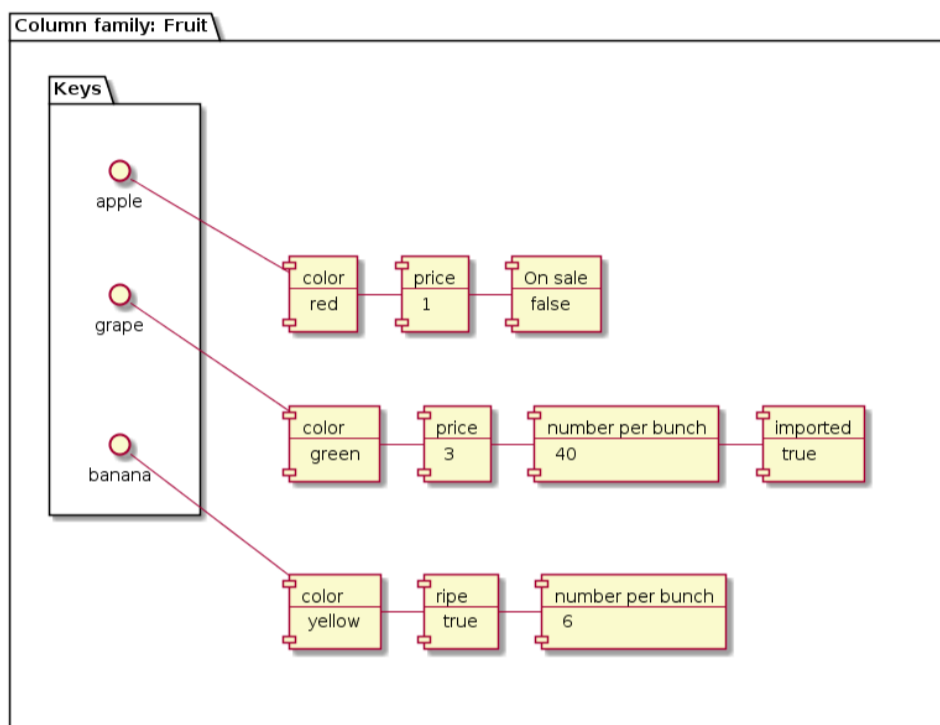
- Для OLTP
- Для OLAP

Назначение

- DB в памяти кэш
- На дисковом хранилище
- В облаке

Хранение

Колоночные хранилища



Для колоночных хранилищ информация записывается в колонки. Колонки между собой не связаны, поэтому можно безболезненно добавлять и удалять новые свойства. Отсутствие четкой заранее заданной схемы позволяет хранить в этих NoSQL БД записи без определенной структуры.

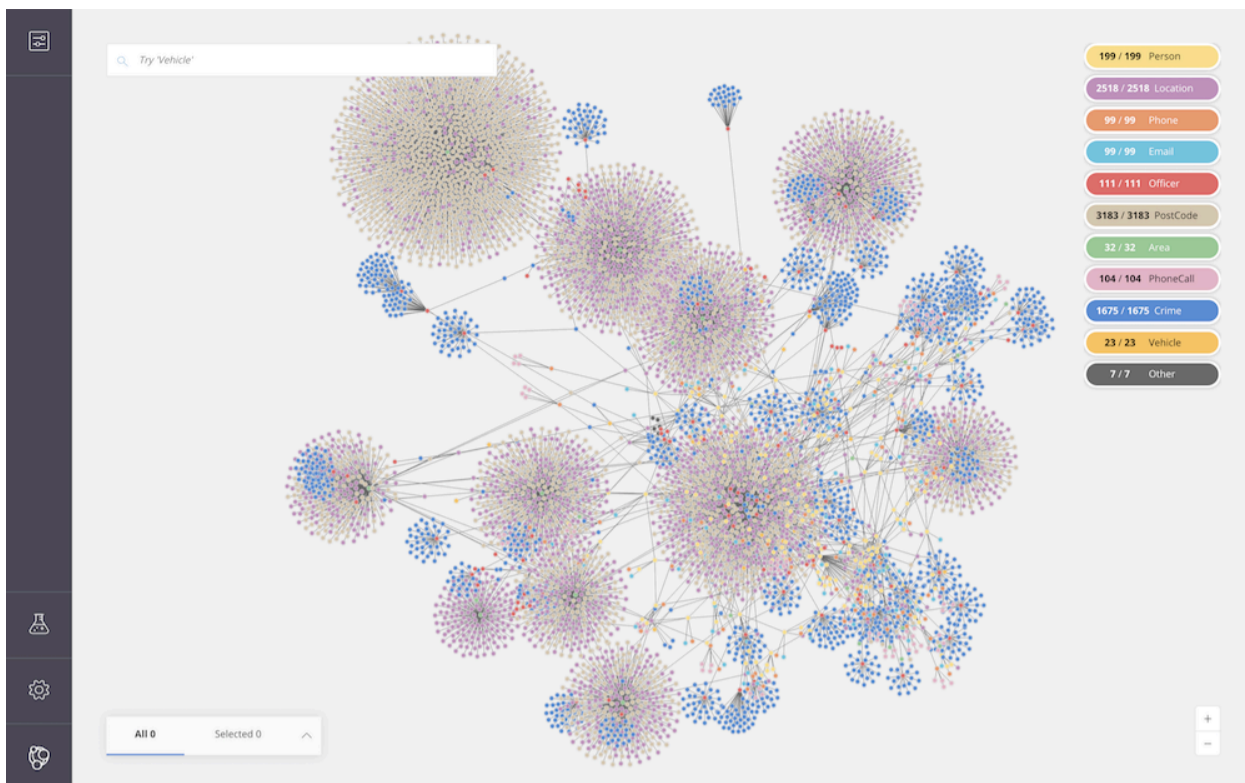
Отличие в том, что реляционные БД при запросе сканируют всю таблицу, а информация из строки извлекается целиком. В колоночных хранилищах поиск ведется по отдельным столбцам и выгружается отдельное значение. Так данные получаются быстрее, и место при хранении экономится.

В колоночных БД имеются структуры – колоночные семейства. Семейства содержат строки, в каждой из которых определен собственный формат. У строк есть собственный уникальный идентификатор для поиска, за которым следуют наборы имен.

Обобщаем:

- Такие БД удобны при работе с приложениями, требующими высокой производительности;
- Данные и метаданные доступны по одному идентификатору;
- Гарантировано размещение всех данных строки в одном кластере, что упрощает сегментацию, масштабирование данных.
- Примеры: ClickHouse, Vertica, Cassandra, HBase.

Графовые базы данных

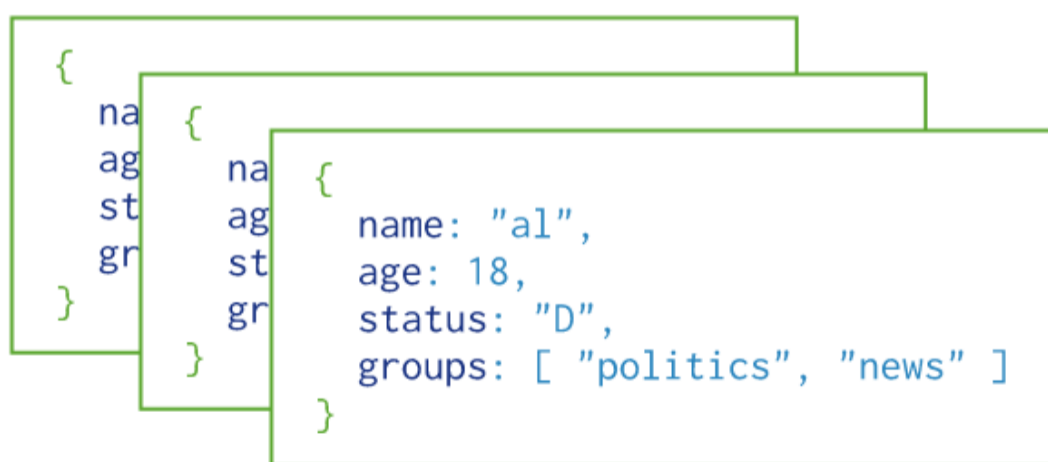


Вместо сопоставления связи с таблицами графовые БД устанавливают связь, используя узлы, ребра и свойства. Графовые БД представляют собой данные в виде узлов, которые могут иметь любое количество связанных с ними свойств. У ребра связи есть начало и конец, направление и тип. Главное отличие от других типов заключается в том, что в реляционных БД соединение между элементами требует дополнительных вычислений. Примеры использования: социальные сети, алгоритмы рекомендаций.

Обобщаем:

- Здесь нет ограничений в типах представления связей.
- Графовые БД не требуют пошагового обхода для перемещения между узлами.
- Они явно отображают связь между типами данных.
- Фокус делается на связях между элементами.
- Примеры таких БД: Neo4j, Giraffe, Tiger-db, Oriana-db.

Документоориентированные хранилища



Документоориентированные хранилища часто представляются в коде как объект-документ. Формат подобен JSON и его легко воспринимать. Характер хранения – слабоструктурированный, иерархический. Документы хранятся в коллекциях. Коллекции группируются логически. Структура от документа к документу может различаться. Документы могут хранить уже привычные массивы, другие документы, JSON'ы.

Документы oriented-модель хороши в проектах, где нужно обрабатывать большой объем данных без четкой структуры: каталоги, социальные сети, платформы с блогами и видео и т.д.

Обобщаем:

- Такие БД не имеют определенной схемы;
- У каждого документа своя внутренняя структура;
- Они подходят для быстрой разработки;
- Можно менять свойства данных, не меняя структуру или сами данные всей базы.
- Одна из самых популярных БД – MongoDB (легкомасштабируемая, отказоустойчивая, данные в ней хранятся в безопасности).
- Примеры: RethinkDB, Amazon DocumentDB.

Хранилища «ключ-значения»

Key	Value
K1	AAA,BBB,CCC
K2	AAA,BBB
K3	AAA,DDD
K4	AAA,2,01/01/2015
K5	3,ZZZ,5623

В хранилищах «ключ-значения» данные хранятся в парах с уникальным идентификатором. Информация может храниться в простом или составном виде.

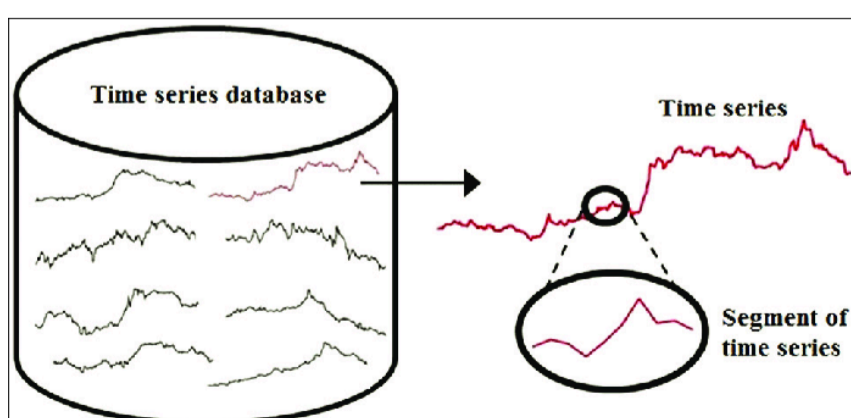
Key-Value часто используется в случае, когда скорость является приоритетом, а данные не слишком сложные. Например, для хранения кэша данных, онлайн списков, обработки истечения сроков действия, разделения сеансов построения рейтинга и т.д.

БД с парами «ключ-значение» поддерживают высокую разделяемость данных и обеспечивают беспрецедентное горизонтальное масштабирование, недостижимое при использовании других видов БД. Примеры: предоставление ключа-объекта в виде JSON, текста или изображения. Вы отправляете ключ и получаете blob-объект – двоичный, большой объект.

Обобщаем:

- Хранилище обеспечивает быстрый и малозатратный доступ;
- Нет жесткой схемы отношений между данными, поэтому в таких БД часто хранят различные типы данных, данные конфигов, информацию о состояниях (часто они предоставлены хэшами или словарями);
- Разработчик отвечает за определение схемы наименования ключей, а также за то, чтобы значение имело соответствующий тип-формат. Пример: Redis (система расположена в оперативной памяти, что ускоряет процесс очищения записи), ETCD, Memcached.

Time-Series DB



БД временных рядов созданы для сбора и управления элементами, меняющимися с течением времени.

Большинство таких БД хранят значения по элементам. Например, можно создать структуру, отслеживающую температуру CPU, внутри таблицы будет храниться составное значение из временной ветки и температуры.

Эти БД ориентированы на запись, предназначены для постоянного потока выходных данных. Производительность зависит от количества отслеживаемых меток, интервала опроса и фактически полезной нагрузки.

Примеры: Cassandra, InfluxDB, Prometheus, TimescaleDB.

Виды по назначению

- Для OLTP используются DynamoDB, Cassandra.
- Для OLAP используются ClickHouse, MapReduce.

Это разделение следует из их структуры и из того, насколько большие данные они умеют обрабатывать и как быстро.

Виды по хранению

Существуют виды NoSQL БД, которые хранят данные в оперативной памяти. Их называют «**In-memory**». Благодаря такому хранению данные преобразовываются в течение нескольких микросекунд, даже если их становится больше.

Примеры: Redis, Amazon (они дают работу практически в режиме реального времени).

Еще один пример – **DynamoDB Accelerator**. БД, хранящиеся на диске, – это почти все, рассмотренные ранее.

Облачное БД – это те базы, которые запускаются на платформах облачных вычислений. Если поставщики облачных VM могут поставить VM для БД, то можно взять и готовую БД как сервис, где обслуживанием будет заниматься поставщик.

Пример первых: Hadoop, MongoDB, Cassandra от Amazon.

Пример вторых: Mongo database servers, Amazon DynamoDB, Cloud Base и т.д.

Как вам урок?



Изучил, далее >

Слёрм ©

[+7 \(495\) 248-05-80](tel:+74952480580)

[Лицензия №ДЛ-1368 от 22.08.2019](#)

[Политика конфиденциальности](#)

[Публичная оферта](#)

