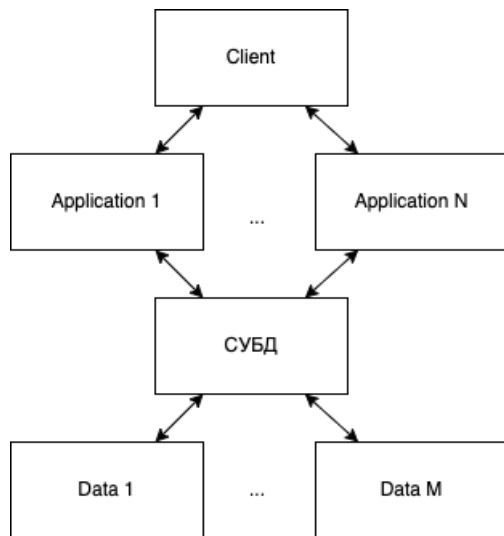


1 Распределенная система БД



1.1 Предпосылки к распределенным БД

- ...
- ...
- Естественная распределенность данных

Распределенная БД – набор, логически связанных между собой разделяемых данных и их описаний, которые физически распределены по нескольким вычислительным узлам

Фрагментирование – разделение одной таблицы на несколько

1.2 Способы фрагментирования

- Горизонтальное – выделение подмножеств строк
Когда есть какой-то временной аспект для данных
- Вертикальное – пытаемся хранить таблицу, не как кортежи со всеми атрибутами, а как несколько кортежей с какими-то атрибутами
Где удобно задать безопасность для ролевой модели (секретная / несекретная области)
- Смешанное – подмножества строк с подмножеством атрибутов
Когда удобно совмещать

Репликация – поддержка синхронизированных физических копий некоторого объекта БД

1.3 Стратегии размещения данных в распределенной системе

- Раздельное (фрагментированное) размещение – БД разбивается на непересекающиеся фрагменты и каждый фрагмент располагается строго на одном узле
- Размещение с полной репликацией – на каждом узле есть полная копия (реплика – синхронизированная копия) всей БД

Большая стоимость хранения

Высокая надежность

Производительность неоднозначна

- Размещение с выборочной репликацией – разделяем БД на фрагменты и для каждого фрагмента пытаемся решить задачи: сколько копий сделать? где копии расположить?

1.4 Принципы прозрачности

- Прозрачность фрагментации
- Прозрачность расположения
- Прозрачность количества реплик
- Прозрачность контроля доступа

Пользователь получивший отказ в доступе не знает существуют ли такие данные или у него нет доступа

1.5 Базы данных

- Гомогенные – используют одинаковую СУБД на всех узлах
- Гетерогенные – используют разные СУБД на всех узлах

1.6 Правила распределенных БД

Идеальная

- Локальная автономность – локальные данные (данные на узле) принадлежат локальным владельцам и локально сопровождаются
- Отсутствие опоры на центральный узел – в системе не должно быть ни одного узла, без которого система не может функционировать
- Непрерывное функционирование – в системе не должна возникать потребность в плановом останове ее функционирования

- Независимость от расположения
- Независимость от фрагментации – получение доступа к данным вне зависимости от их фрагментации
- Независимость от репликации – пользователь не должен нуждаться в сведениях о наличии реплик
- Обработка распределенных запросов – обработать любой запрос вне зависимости на скольких узлах расположены требуемые объекты данных
- Обработка распределенных транзакций – система должна поддерживать транзакции с данными, расположенными более чем на одном узле
- Независимость от типа оборудования
- Независимость от сетевой архитектуры
- Независимость от ОС
- Независимость от типа СУБД

1.7 Задачи обработчика распределенных запросов

- В каком фрагменте расположены нужные мне сведения?
- К какой копии фрагмента обращаться?
- Где расположить временные структуры?

1.8 Преимущества распределенных систем

- Нативно отображает структуру организации
- Разделяемость и локальная автономность
- Повышение доступности и надежности
- Повышение производительности
- Модульность системы - позволяет распараллелить обработку и обслуживание

1.9 Недостатки распределенных систем

- Повышение сложности на всех уровнях (архитектура, алгоритмы и тп)
- Увеличение стоимости владения
- Проблема защиты

- Усложнение контроля за целостностью данных
- Отсутствие стандартов