

Базы данных, лекция 8

@mikhirurg

April 2020

1 Надёжность хранения данных

Необходимо обеспечить доступность данных, мочь в любой момент времени обратиться к данным и получить их за конечное время отклика.

1.1 Физическая сохранность данных

- Дисковые массивы с избыточностью (RAID-массив, СХД)
- Резарвирование аппаратных компонентов и сетей доступа к данным
- Центры обработки данных
Парковка автомобилей сотрудников должна быть в 600 метров от ЦОД, их строят в сейсмически нейтральных зонах. Бывают мобильные ЦОД'ы.

1.2 Логическая сохранность данных

- Сложно идентифицировать нарушения логич. сохранности.
- Невозможно обеспечить целостность данных в каждый момент времени в процессе выполнения задач по их изменению. Логически согласование состояния БД.

Пример: Снятие средств с одного счёта и передача на другой. Между двумя действиями произошёл сбой. Деньги исчезли. Система пришла в логически неверное состояние.

1.3 Транзакция

Транзакция - это последовательность действий с БД, в которой либо все действия были успешны, либо не выполняется ни одно из них.

Два результата завершения: (Commit; Rollback)

Свойство транзакции:

Похоже на
Git-
репозиторий

- Атомарность
- Согласованность
- Изоляция

Если запущено несколько конкурирующих транзакций, результат выполнения каждой из них должен быть скрыт от другой

- Долговечность

Необходимо зафиксировать выполнение транзакции, чтобы последующие возможные сбои не повлияли на состояние данных.

Пример: Перевод денег между абонентами сотовой сети

Списать суммы у одного пользователя двумя разными транзакциями

(В результате одной из транзакций он ушел в минус)

1.4 Проблемы конкурирующих транзакция

- Проблема потерянного состояния
(пример про абонента) каждая из транзакций считала баланс, поняла что нужно снять 100р. и в итоге записала 100 ему на счёт.
В итоге абонент остался с несуществующими 100 рублями.
- Проблема "грязного чтения"
Читается кортеж другой транзакции, который должен откатиться. То есть транзакция считала неверные данные.
- Проблема неповторяемого чтения
В первый раз, когда было считывание, сумма была одна. А при записи итоговая сумма оказалась другая. Запись в несогласованное состояние.
- Проблема чтения фантомов

1.5 Блокировки

Можно разделить на явные и неявные блокировки

Объект блокировки:

- Строка
- Столбец
- База данных

Нужно выбирать правильный объект для блокировки

Бывают монопольные и коллективные блокировки.

Удержание транзакций на время блокировки

Таким образом мы решаем проблемы конкурирующих транзакций.

1.6 Уровни изоляции транзакций

- Незавершённое чтение
если транзакция начала изменять данные, другая не должна менять их
- Завершённое чтение
Разрешаем чтение если транзакция завершилась
- Воспроизводимое чтение
Если одна транзакция начала читать данные, нужно запретить другим транзакциям менять зависимые данные
- Сериализуемость
Нужно обеспечить то, что появление новых кортежей невозможно, пока не будет снята соответствующая блокировка.
Блокируются объекты, подпадающие под определённое условие.

Таким образом мы обеспечиваем изоляцию транзакций.

1.7 Журналирование транзакций

В некоторый журнал записывается план транзакции, транзакция выполняется и после выполнения транзакция стирается из журнала (или её часть). Если транзакция не завершилась, по журналу производится Rollback до контрольной точки или до транзакции (зависит от организации самой транзакции).

(Обеспечиваем долговечность транзакции).