

Транзакции это хорошо и надёжно, но что если у нас в базе данных не один пользователь а 10.000.000? Как транзакции будут выполняться и не влиять друг на друга? Возникает параллелизм.

Параллелизм – это мощный инструмент, который позволяет увеличить скорость обработки запросов в системе, путем одновременного выполнения множества операций. Однако использование параллелизма может привести к сложным проблемам согласованности данных, особенно в контексте транзакций. Среди этих проблем можно выделить следующие:

1. Потерянное обновление (Lost Update): это происходит, когда две или более транзакции читают одну и ту же запись, затем изменяют ее и записывают обратно. Если вторая транзакция записывает свои данные после первой, то обновления, внесенные первой транзакцией, могут быть потеряны. Для борьбы с этой проблемой используются механизмы блокировки или многоверсионности, которые гарантируют, что только одна транзакция в определенный момент времени может работать с определенной записью.
2. Грязное чтение (Dirty Read): это происходит, когда одна транзакция читает данные, которые были изменены другой транзакцией, но еще не зафиксированы (коммитнуты). Если изменяющая транзакция впоследствии откатывается (отменяет свои изменения), то читающая транзакция может обнаружить, что она работала с "грязными" (некорректными) данными. Для предотвращения грязного чтения уровень изоляции транзакций должен быть установлен по крайней мере на уровень "read committed".
3. Неповторяющееся чтение (Non-Repeatable Read): это происходит, когда в рамках одной транзакции одни и те же данные считываются несколько раз, и во время этих чтений другая транзакция меняет эти данные. В результате одна и та же транзакция может получить разные результаты для одного и того же запроса, что может привести к неконсистентным или неправильным результатам. Для предотвращения неповторяющихся чтений уровень изоляции транзакций должен быть установлен по крайней мере на уровень "repeatable read".
4. Фантомное чтение (Phantom Read): это вариант неповторяющегося чтения, который происходит, когда в рамках одной транзакции выполняется запрос, выбирающий набор строк, и другая транзакция в это время добавляет или удаляет строки, которые подходят под условие этого запроса. В результате первая транзакция может обнаружить "фантомные" строки – строки, которые не были там, когда она в первый раз выполнила запрос. Чтобы предотвратить фантомное чтение, уровень изоляции транзакций должен быть установлен на уровень "serializable", что гарантирует, что каждая транзакция будет исполняться, как будто она единственная в системе, обеспечивая таким образом максимальную изоляцию и согласованность данных.

Управление этими проблемами основывается на использовании правильного уровня изоляции транзакций, а также на тщательной координации и синхронизации между транзакциями в системе.