

3. Что нельзя делать с таблицами

- TRUNCATE/DROP PARTITION – хорошо, ALTER TABLE ... DELETE/UPDATE ... – очень плохо
- Нельзя запускать сотни/тысячи вставок (INSERT) в секунду
- Нельзя делать альтеры на продовых процессах
 - Пример, как можно сломать таблицу. Так делать не нужно!
- Нельзя использовать конструкцию CREATE TABLE ... ON CLUSTER ... AS SELECT

TRUNCATE/DROP PARTITION – хорошо, ALTER TABLE ... DELETE /UPDATE ... – очень плохо

Удалить партицию – легковесная операция, для этого ClickHouse достаточно пометить файлы партиции как удаленные.

В отличие от OLTP баз, например PostgreSQL, удаление/обновление в ClickHouse – это тяжелая операция. Для выполнения `ALTER TABLE ... DELETE/UPDATE` ClickHouse **полностью пересобирает файлы**, в которых находились удаленные /измененные строки.

Если нужно стереть всю таблицу, вместо `ALTER TABLE ... DELETE WHERE 1=1` используйте `TRUNCATE TABLE ...`



`ALTER TABLE ... DELETE/UPDATE ...` не предназначены для регулярных выполнений. Вместо этого используйте stage-таблицы в своих ETL процессах. [Подробнее.](#)

Нельзя запускать сотни/тысячи вставок (INSERT) в секунду

Оптимальный темп вставки данных – 10-500 тысяч строк раз в 1-2 секунды. [Подробнее.](#)

В пайплайнах, например в задачах Airflow, стоит делать `INSERT` единожды.

Нельзя делать альтеры на продовых процессах

ClickHouse хорошо работает с редким чтением и вставкой огромного количества строк, но плохо работает с изменением /удалением строк (мутациями) и частой вставкой маленьких объемов.

Более того, **это опасно**: после того как операция выполнялась, продолжают идти процессы мерджей, догоняются реплики и т.д. Если в этот момент по ошибке запустить еще один `ALTER TABLE ... ON CLUSTER`, то может произойти сбой и полная остановка всех DDL операций, включая операций ваших коллег.

Пример, как можно сломать таблицу. Так делать не нужно!

Без минимальной задержки были запущены запросы

1. `TRUNCATE table ON CLUSTER shard_group_old`
2. `INSERT INTO table`
3. `ALTER TABLE table ON CLUSTER shard_group_old ADD COLUMN IF NOT EXISTS current_time`
4. `ALTER TABLE table ON CLUSTER shard_group_old UPDATE update_time = current_time WHERE 1 = 1`
5. `ALTER TABLE table ON CLUSTER shard_group_old DROP COLUMN IF EXISTS current_time` – упало с ошибкой, т.к. какие-то мутации не могут выполняться, т.к. отсутствует колонка `current_time` в каких-то партах на какой-то ноде

Не используйте апдейты!

Самый безопасный способ изменения существующих данных – создание правильной таблицы в `sandbox` схеме и последующая переливка, как есть, в целевую.

Нельзя использовать конструкцию CREATE TABLE ... ON CLUSTER ... AS SELECT

[Подробнее про CREATE TABLE ... ON CLUSTER ... AS SELECT.](#)