

1. При выполнении 1 операции INSERT 40% времени уходит на поиск по кластерному индексу и его обновление, 30% на вычисления и всего 10% непосредственно на вставку строк в таблицу Payment. В среднем данный запрос выполняется 10 секунд.

При вставке 500 строк отдельными операциями, прошло больше 40 минут и транзакции закрылись аварийно(видно не хватает мощности наших стареньких ноутов).

При вставке данных не отдельными строками, а по нескольку строк за раз, при помощи команды BULK INSERT, время выполнения сокращается до 1.3 секунды при вставке 10 строк. При вставке 500 строк за раз запрос исполняется в среднем около 2 секунд. Больше всего времени при обоих запросах тратится на поиск по кластерному индексу(около 30%) и его обновление .

2. После оценки времени выполнения запросов стало понятно, что лучше всего накапливать запросы в течение какого-то количества времени, а затем вставлять блоками в целевую бд. На практике надо высчитывать среднее количество запросов, которые накапливаются за определенный промежуток времени(например, за 5 часов или сутки и т.д.) и выполнять их одним блоком.

В качестве реализации предлагается записывать данные в файл .csv, из которого база считывает данные. В данном случае предполагается, что оператор будет вести работу с базой при помощи веб-приложения. Данные, отправляемые с сервиса, будут иметь формат JSON, который удобно хранить в .csv формате. В данном случае первичная проверка правильности ввода данных будет проводиться на фронтэнде.

3. Так как мы используем вариант отложенных вычислений, пользователь не сразу может увидеть изменение баланса, что вызывает неудобство. Например, при считывании данных, аналитик может получить старые данные. С другой стороны, если задержка по времени будет стабильной, и не очень долгой, то пользователи привыкнут и не будут беспокоиться о незавершенности или какой-либо ошибке во время операции.