

Рассчитанные цифры нагрузки позволяют нам добавить **конкретные количественные значения к нефункциональным требованиям**, которые мы ранее определили. Это значит, что мы можем перейти от общих описаний качества, таких как "система должна быть быстрой" или "система должна быть надежной", к более конкретным и измеримым нефункциональным требованиям к качеству, таким как "система должна обрабатывать X транзакций в секунду" или "система должна иметь время простоя не более Y часов в год".

Это не только улучшает нашу способность проектировать и реализовать систему, которая соответствует требованиям, но также позволяет нам более точно оценить ресурсы, необходимые для поддержки системы, и предсказать ее поведение в реальных условиях.

Так что давайте приступим к добавлению этих конкретных цифр к нашим нефункциональным требованиям и продолжим наш путь к созданию высококачественного решения.

- Производительность (Performance):
  - Задержка ответа чат-бота (latency в секундах): Бизнес-требование заключается в том, чтобы предоставить клиентам быстрый и эффективный сервис. Соответственно, важной метрикой будет задержка времени ответа, которое чат-бот тратит на ответ клиенту. Это можно измерить, отслеживая время между получением вопроса и отправкой ответа. Задержка - менее 3 секунд (что очень хорошо, учитывая что время ответа оператора до 3х минут)
- Надёжность - Отказоустойчивость(Fault tolerance):
  - Отказ какого-то компонента системы (в штуках): Система должна быть способна продолжить работу без сбоев, даже в случае отказа одного или нескольких из её компонентов. В случае возникновения ошибки, система должна автоматически перенаправить запросы от отказавшего компонента к резервному. Время восстановления после отказа не должно превышать 5 минут.  
На основе наших метрик мы можем предположить что время восстановления приемлемое.
- Доступность (Availability):
  - Время простоя (измеряют в процентах 99,(количество девяток, чем больше, тем ближе к 100% доступности)): Бизнес-требует, чтобы система была доступна в любое время для поддержки клиентов. Метрика здесь может быть общее время простоя системы в течение определенного периода (например, месяц, день, год).

Можем установить время доступности в 99.9. Так как у нас мало пользователей - нет необходимости закладывать больше потому что вероятность недоступности будет очень низкая.

- Масштабируемость (Scalability):

- Возможность добавления пользователей (нагрузка в RPS - кол-во запросов в секунду): Из вводных следует, что клиентов может быть много. Поэтому одной из метрик масштабируемости будет способность системы поддерживать увеличение числа одновременных пользователей без ущерба для производительности.

Судя по нашим метрикам - у нас не так много нагрузки, поэтому запишем 40 RPS и зафиксируем, что мы можем выдерживать и в монолитной архитектуре.

- Безопасность (Security):

- Нарушения защиты данных: Бизнес-требуется, чтобы персональные данные клиентов были защищены (обеспечение шифрования и защиты данных от несанкционированного доступа).

Тут можно уже зафиксировать конкретные данные, которые допустим были обсуждены с отделом безопасности - Система должна обеспечивать строгую защиту персональных данных клиентов в соответствии с Федеральным законом "О персональных данных" №152-ФЗ Российской Федерации

Отлично! теперь у нас есть расчеты по нагрузке и чёткие нефункциональные атрибуты качества.