

Стратегия Масштабирования Сервиса

Если сервис не очень большой, можно воспользоваться cloud services, например AWS от Амазона. Но этот вариант не подходит, то есть еще два пути по которым можно пойти - масштабирование оборудования и программного обеспечения.

Оборудования можно усовершенствовать путем горизонтального и/или вертикального масштабирования. При горизонтальном масштабировании добавляется количество серверов, а при вертикальном их вычислительная мощность. Иметь один большой но очень мощный сервер ненадежно, а иметь много маленьких обернется в головную боль системному администратору, по этому на практике лучше комбинировать два метода масштабирования и пользоваться плюсами обоих способов.

Но в основном приходится иметь дело с масштабированием программного обеспечения. Большая часть сервисов состоит из трех частей: сервера, приложения и базы данных. Масштабировать можно все три компонента. Далее пойдет перечисление некоторых способов оптимизации:

1. Убедиться что данные хранятся правильно, а это значит что они должны лежать в базе данных (SQL/ NoSQL). Это важно потому что в базе данных скорость поиска нужной информации во много раз быстрее чем поиска на обычном жестком диске;
2. Настроить кэш сервер чтобы ускорить часто используемые запросы;
3. Минимизировать количество запросов на базу данных и сервер;
4. Использовать балансировку нагрузки, если имеется много машин;
5. Использовать оптимальные алгоритмы в приложении;

Масштабирование это дело хорошее и нужное, но никогда не надо этим заниматься, если в этом нет необходимости. Как говорил Дональд Кнут: "Преждевременная оптимизация есть корень всех зол".