

Введение

В этом уроке мы научимся эффективно работать с файлами, что является не такой тривиальной задачей, как может показаться на первый взгляд.

С маленькими файлами работать просто. Они целиком помещаются в памяти программы, ими легко оперировать, их можно быстро передавать. Проблемы начинаются при работе с большими файлами.

Такие файлы уже нельзя так просто хранить целиком в памяти веб-сервера при обработке. С ними нужно работать частями.

Streaming

Подход, который позволяет эффективно работать с файлами. Основная идея в том, что в большинстве операций с большими файлами веб-сервер является *проху* между клиентом и файловым хранилищем. Поэтому нам не нужно хранить весь файл целиком. Достаточно оперировать в один момент времени небольшими частями этого файла.

Эти части в дальнейшем мы будем называть «чанки» (chunk).

Примеры, когда streaming подходит, а когда — нет

- вы делаете lms metaclass, и в материалах на платформе есть достаточно длительные видеозаписи. В качестве хранилища записей используется vimeo или S3-хранилище. В этом случае оправдан подход с чанками, потому что видео может быть тяжелым. Мы не обрабатываем его на лету, а просто проксируем в хранилище
- вы делаете социальную сеть, и нужно сделать загрузку аватарок пользователей. В этом случае картинки относительно маленькие. К тому же велика вероятность, что картинку нужно будет как-то обработать, а для этого нужен весь файл. Поэтому подход с чанками использовать не стоит.