

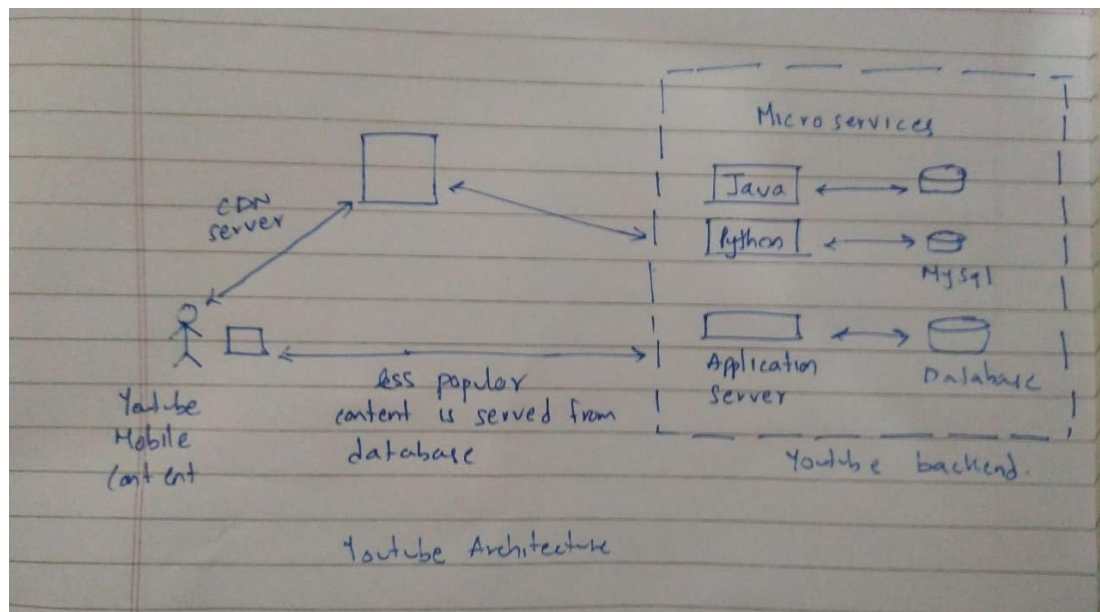
ISE Question 1

Name: Ritik Singh

UID: 2018130052

Roll Number: 58

A) Popular videos are served from the CDN & moderately, lesser-played videos are fetched from the database. Every video, at the time of the upload, is given a unique identifier and is processed by a batch job that runs several automated processes such as generating thumbnails, metadata, video transcripts, encoding, setting the monetization status and soon.



The videos are stored in the hard drives in warehouse-scale Google data centers. The data is managed by the Google File System & BigTable. GFS Google File System is a distributed file system developed by Google to manage large scale data in a distributed environment.

B) All the videos that are uploaded to YouTube are first transcoded into multiple different formats and resolutions set by the platform. The video during the transcoding process is broken down into segments and converted into multiple different resolutions.

The processing of multiple segments is spread across multiple machines to parallelize the process thus increasing the throughput. If a video goes viral, it is subject to another round of video compression. This second round of compression ensures the same visual quality of the video at a much smaller size.

When encoding videos YouTube chooses a bitrate within the limits that the codec allows. A video with a high bitrate has better quality but there is a sweet spot beyond which even on increasing the bitrate there is not a significant visual improvement in the video quality though the video size increases in the process.

YouTube also takes into account the playback statistics of the video player on the client's device like how often the player switches to a lower resolution with respect to the streaming bandwidth available on the client & so on. It then predicts if a better resolution content can be pushed to the user, being aware of his streaming bandwidth limit. This performance estimation enabled YouTube to cut down on the bandwidth, increasing user engagement on the platform simultaneously. For encoding its videos YouTube uses VP9 – an open-source codec that compresses videos with HD & 4K quality at half the bandwidth used by other codes. Once the videos are transcoded and stored in the database, they are spread across the cache network of the platform. When the user requests a video, the platform checks the viewer's device type, screen size, processing capability, the network bandwidth and then delivers the fitting video version in real-time from the nearest Edge location.