

System design of Video streaming platform like Hotstar

Requirements

In-Scope

- 1.Content creators should be able to upload their media.
- 2.Viewers should be able to view content with varying screen sizes and Internet speed.
- 3.Platform should have search functionality.

Out-of-scope

- 1.How the billing or subscription works
- 2.How Personalized recommendation works

Onboard a movie/video:

Stateless Services: All services created will be stateless service. These services are designed such that any service instance can serve any request in a timely fashion and so if a server fails it's not a big deal. In the failure case requests can be routed to another service instance and we can automatically spin up a new node to replace it.

Load Balancer: All the requests from client to any service passes through Load Balancer. Load Balancer will distribute incoming requests to any available server capable of fulfilling them. We can use **consistent hashing** here to distribute our load across servers.

$serverIndex = hash(key) \% N$, where N is the size of the server pool.

Video can be of any size. Say a movie of 1.5 hour may be of size 50 GB or 2TB when uploaded.

Streaming 50 GB video directly to end users having different devices and resolutions may not give a good user experience. They may face buffering while playing videos. Each device has a video format that looks best on that particular device.

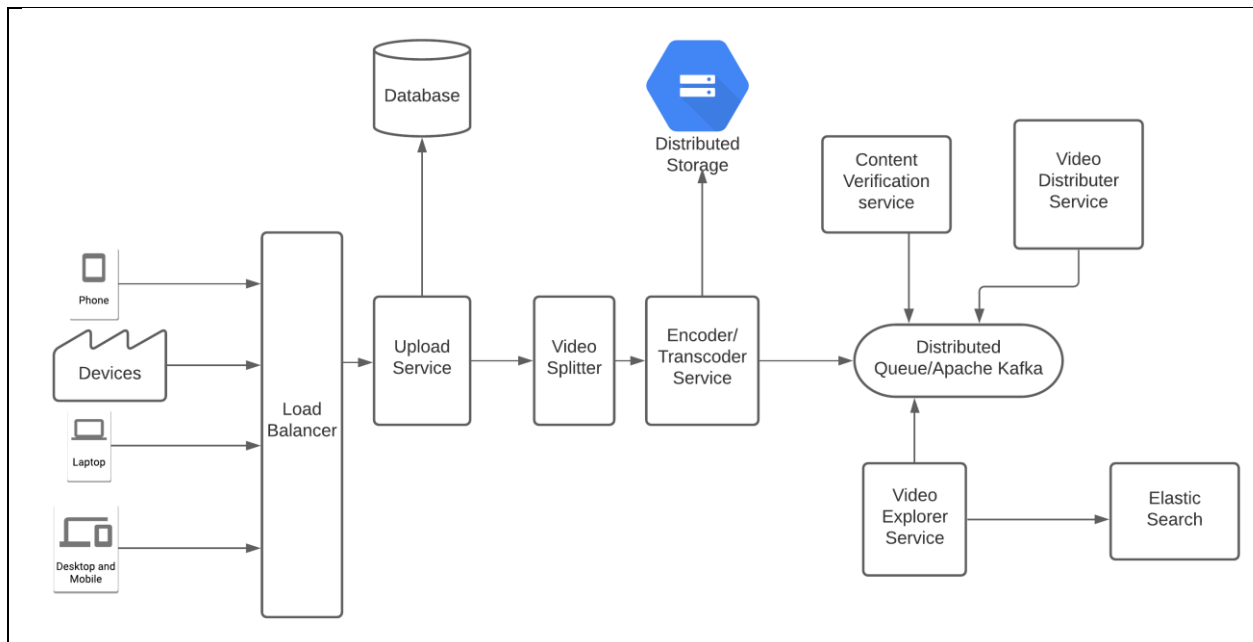
To give users a good experience across all devices, we can split the video using **Video splitter** before storing it in Distributed Storage. Video is stored in variety of resolutions (1080p,720p,480p,360p etc.) and codecs (mp4, flv etc.).

No of formats = No of Codecs * No of Resolutions

= 1200 (approx.)

Hotstar converts these videos in many formats and creates files optimized for different network speeds. Converting the video into a format that works best of a particular device is called transcoding or encoding. Transcoding is done by **Encoder Service**.

Transcoding is the process that converts a video file from one format to another, to make videos viewable across different platforms and devices.



Content creator will upload the video. Load balancer will transfer the request to appropriate Upload service. Upload service will send the video for splitting and

transcoding before being stored in distributed storage. Distributed storage can be a cheap storage like amazon s3 or a distributed file system kind of storage.

Upload service will also put a metadata in the database which can be a relational database like MySQL. It will have a table which will have what is the video that we are uploading, who is the author of the video and what is the url in the distributed file storage.

Video metadata Data Model

Video Id(Primary Key)
Title
description
File_storage_url
Content_creator

Content Verification Service will verify what is the content of the video. Now to verify that, there will be some algorithm and some rules that the video should not contain any offensive kind of language or abusive language or it should not contain anything which is really violent. For each of the chunk, if something like that is found so that video will be blocked and some notification will be sent to the client saying that your video is not up to the mark to be uploaded. If video content is good, then it will say Video is fine.

Video Explorer Service is important is because it will be able to create some tags for the video and it will be able to pass on the metadata to the elastic search based on which your video will become searchable. Whenever a client has uploaded a video so client will be uploading the video along with some detail (detail will be a small description of the video say cast of movies, director, genre of the movie)

It will pass all those things onto elastic search. If any user will search the movie based on genre, elastic search will bring list of all videos based on that genre(say comedy)

Elastic search will have table to store video names based on (genre)tags.

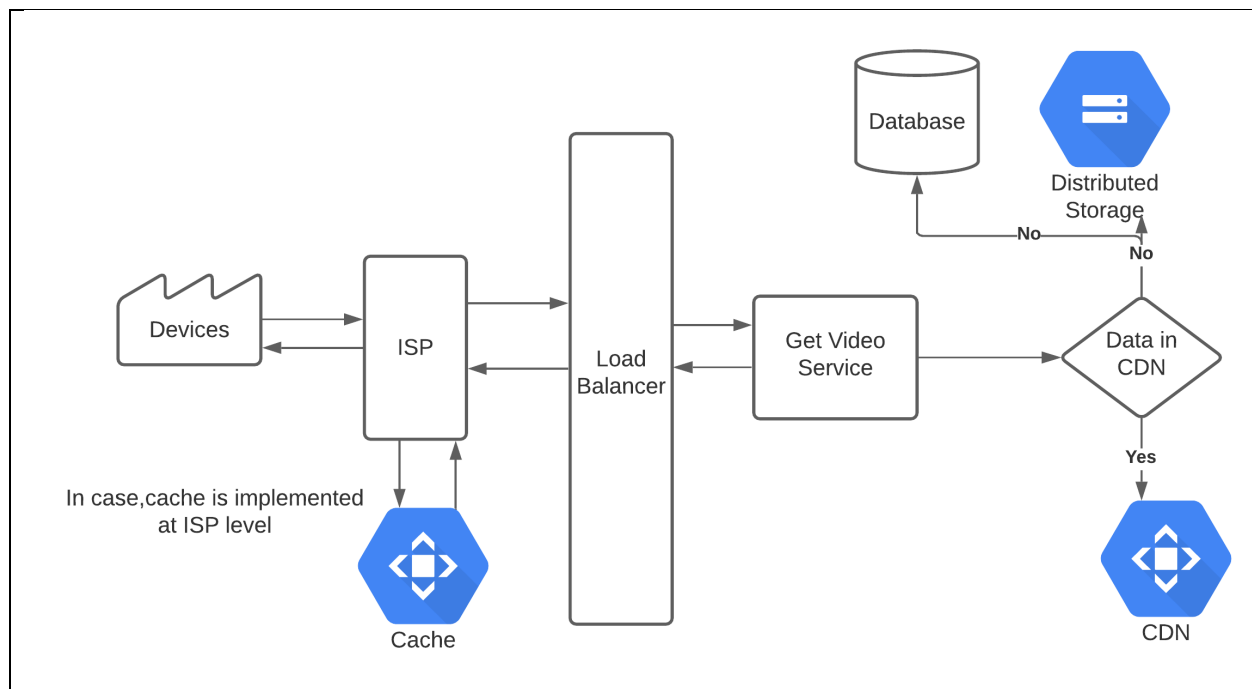
Once elastic search will save the video and its related information, it will send event back to the queue that Video is ready to be distributed.

Video Exploration Service -It will be a costly operation if Distributed storage is situated in North America and user from India is requesting for the video (video files will be large and has to be streamed seamlessly). Delivery of video to users

will be done by cache server. It will distribute videos across cache servers in during an offline hour or less loaded hour. Here it will be stored in CDNs.

CDN - A content delivery network (CDN) is a system of distributed servers (network) that deliver pages and other Web content to a user, based on the geographic locations of the user, the origin of the webpage and the content delivery server.

Viewing a video-



Get Video Service-When a user selects a particular video requests for a video from any device say smartphone, TV or laptop, request will go to ISP (Internet service provider). Request will be like

URL/video_name(<https://www.hotstar.com/movie/movieName>).

ISPs will have table that contains the mapping of IP address to the url name.

id	url	Ip_address
1	https://www.hotstar.com	http://10.1.1.12

ISP will convert the URL to corresponding IP address of Hotstar server (in our case)

When a video is requested, we will pick up the first chunk of that video (we have split the video into multiple chunks) and that metadata will be available here in our database and the actual path will be present or the actual video will be in the distributed storage so that part of the or that chunk of the video will be sent back to the requested device. Distributed storage may be at some location say in North America. If User is requesting video from India, there will be lag in receiving the video and there can be many users like that.

To avoid that, we put videos in cache servers in each of the zone locally. We can store our data in CDNs or use cache servers at ISP level itself.

We can smartly use CDNs here. Suppose a request is going from CDN from India, what is more viewed in India can be stored at CDNs in India

In case a video is requested which is not there in cache, then first chunk of video will be forwarded from Distributed storage and subsequent chunks can be transferred from Distributed storage to cache server. User will get next chunk of video will be served from cache server and will get a seamless service.

Client-side application can also sense the internet connection and request the video in particular format so that buffering can be avoided and video can be played in appropriate format.