High Level Design

Hotstar

Tanmay Kacker

tanmay.kacker@scaler.com

Contents

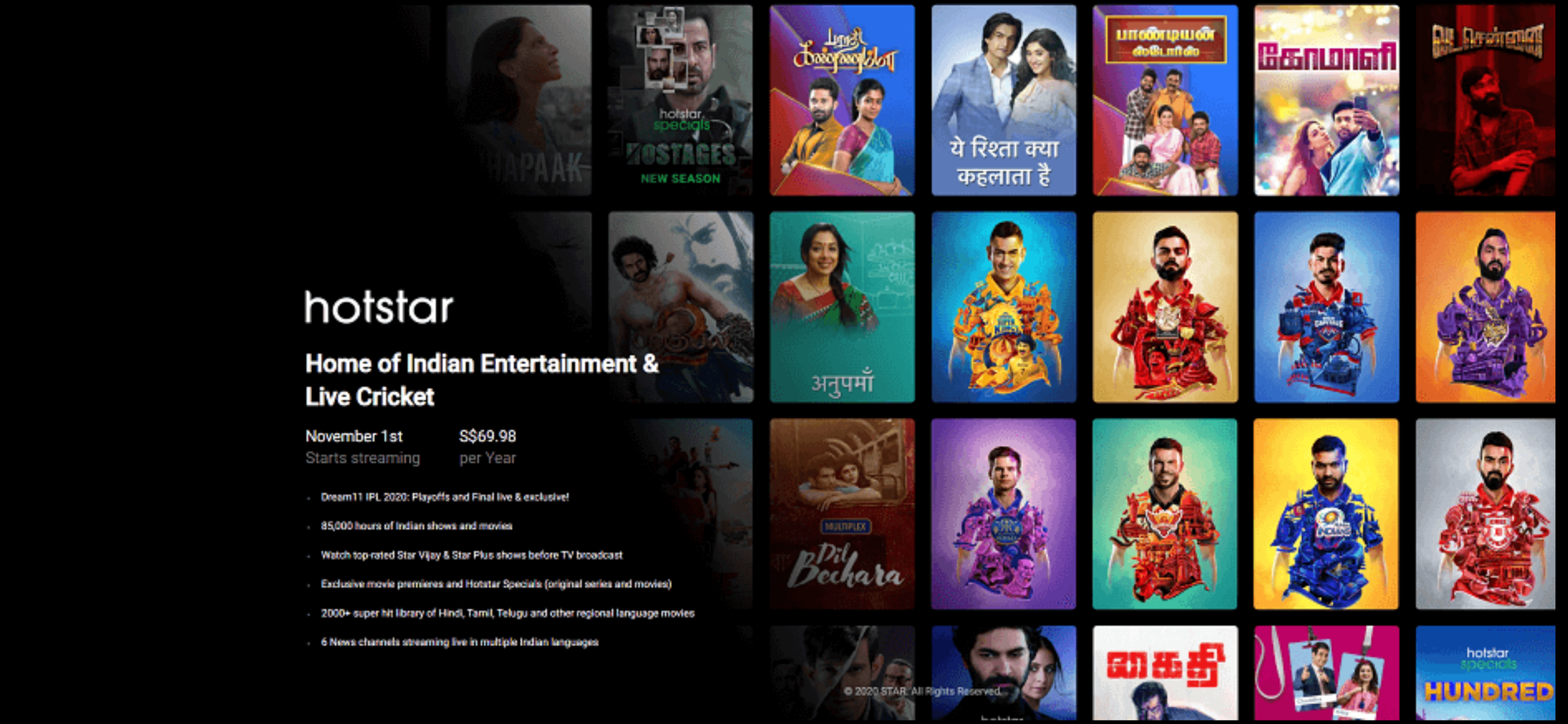
- Problem Statement & Analysis
- Media Processing pipeline
- Data Ingestion architecture
- Video Streaming architecture

Case study

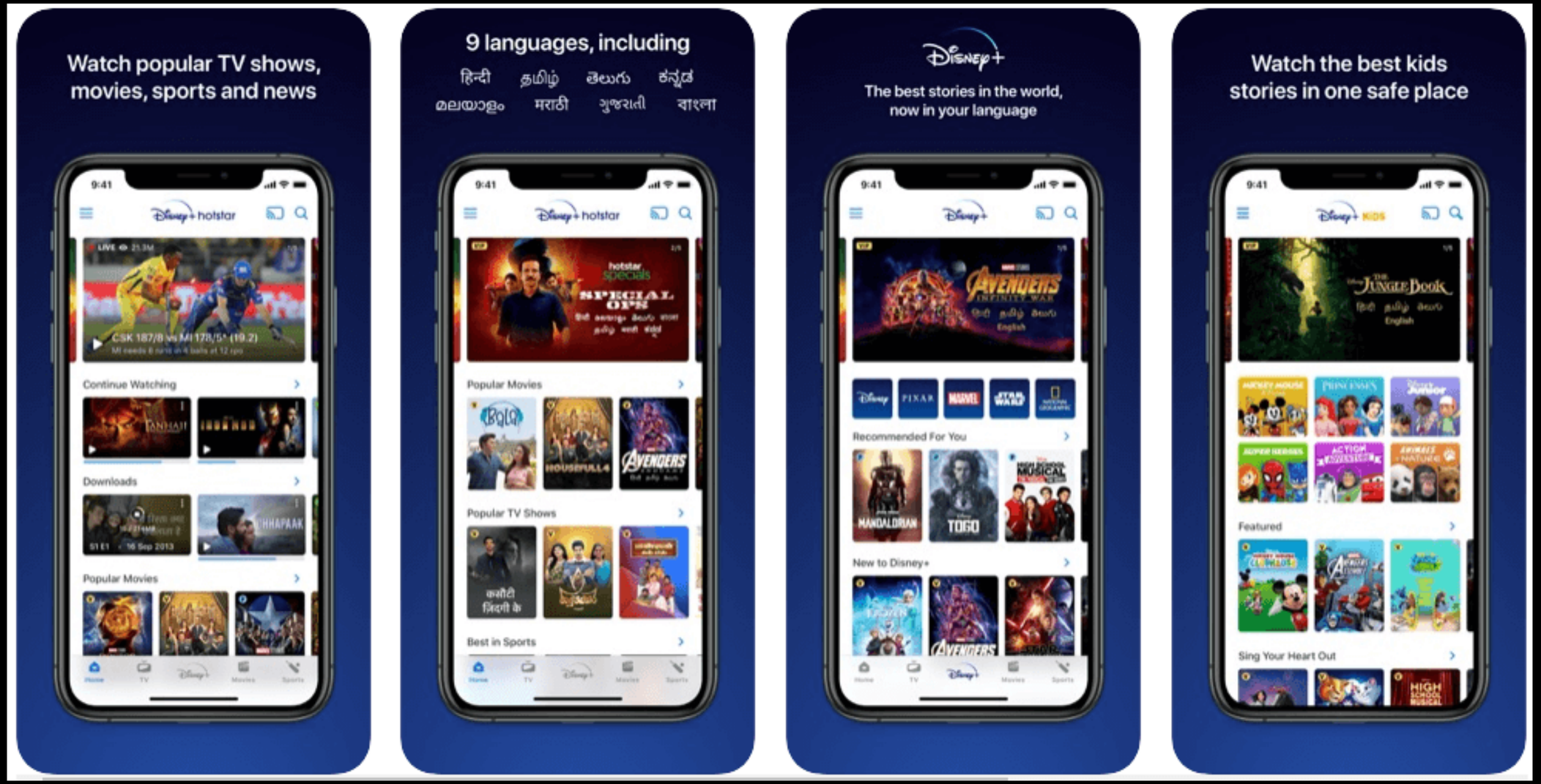
hotstar

- Video on-demand
- OTT streaming platform
- Live video broadcasting

hotstar


hotstar
Home of Indian Entertainment &
Live Cricket
November 1st
Starts streaming
S$69.98
per Year
Dream11 IPL 2020: Playoffs and Final live & exclusive!
85,000 hours of Indian shows and movies
Watch top-rated Star Vijay & Star Plus shows before TV broadcast
Exclusive movie premieres and Hotstar Specials (original series and movies)
2000+ super hit library of Hindi, Tamil, Telugu and other regional language movies
6 News channels streaming live in multiple Indian languages
hotstar specials
HOSTAGES
NEW SEASON
ये रिश्ता क्या कहलाता है
अनुपमाँ
Dil Bechara
HUNDRED
© 2020 STAR. All Rights Reserved.

hotstar


Watch popular TV shows, movies, sports and news
Continue Watching
Downloads
Popular Movies
9 languages, including
हिन्दी தமிழ் తెలుగు ಕನ್ನಡ
മലയാളം मराठी ગુજરાતી বাংলা
Popular Movies
Popular TV Shows
Best in Sports
The best stories in the world, now in your language
Recommended For You
New to Disney+
Watch the best kids stories in one safe place
Featured
Sing Your Heart Out

hotstar

Numbers

- 300 million Monthly Active Users (MAU)
- 100 million Daily Active Users (DAU)
- 46.4 million subscribers
- 100,000 hours of content
- Largest concurrent audience of 19 million users

hotstar

Requirements

Viewer requirements

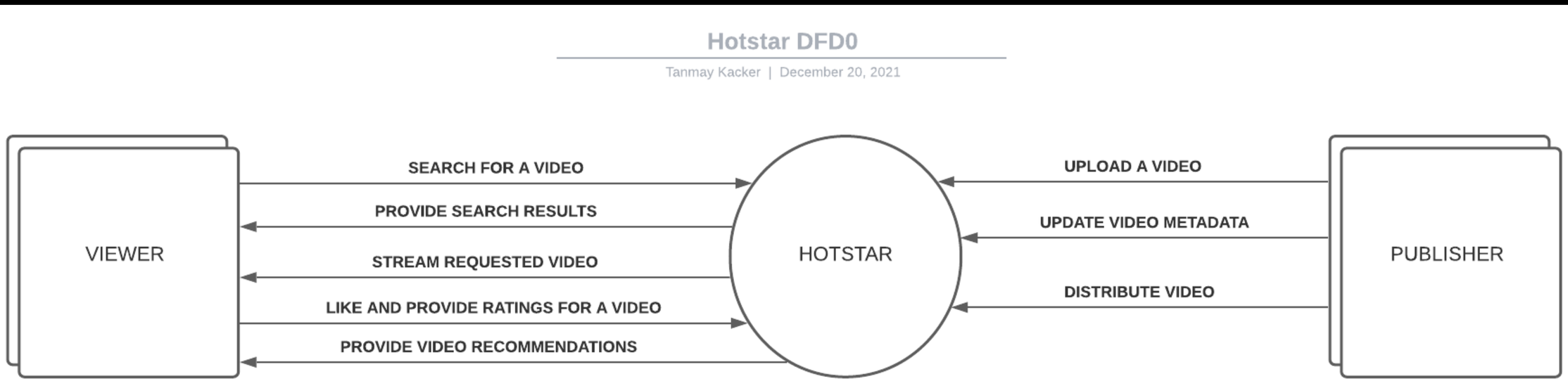
- Search for a video
- Play a video
- Like and rate a video
- Get recommendations

Publisher requirements

- Upload a video
- Distribute a video

hotstar

Data flow



hotstar

Scale estimation

Data Ingestion

- 1 Million videos
- 25% Year on Year growth
- Size of videos - 100 MB to 10 GB

hotstar

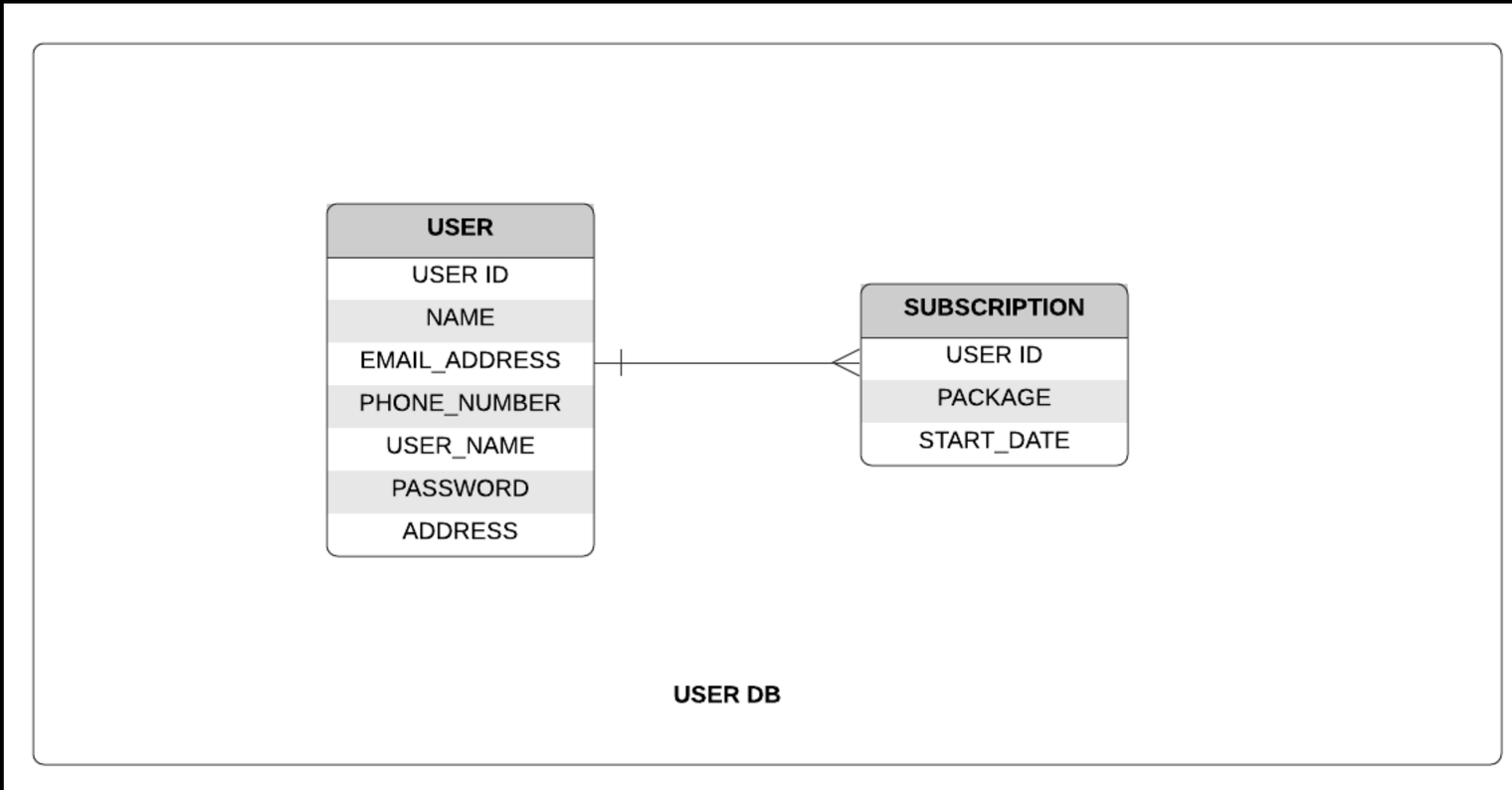
Scale estimation

Video Playback

- 20 million concurrent users
- RDS stores metadata for each video
- 1 KB per video

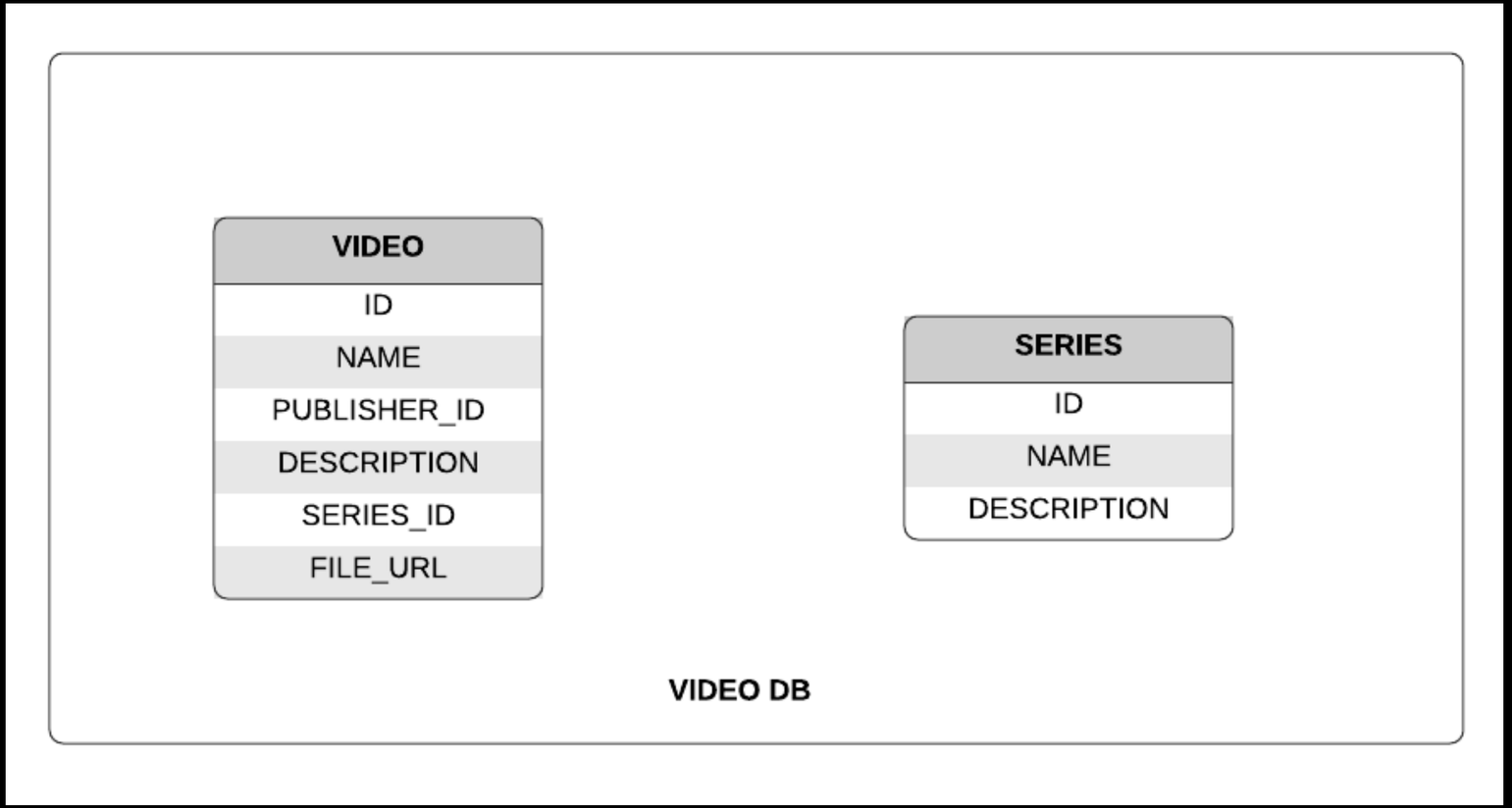
hotstar

Data Models



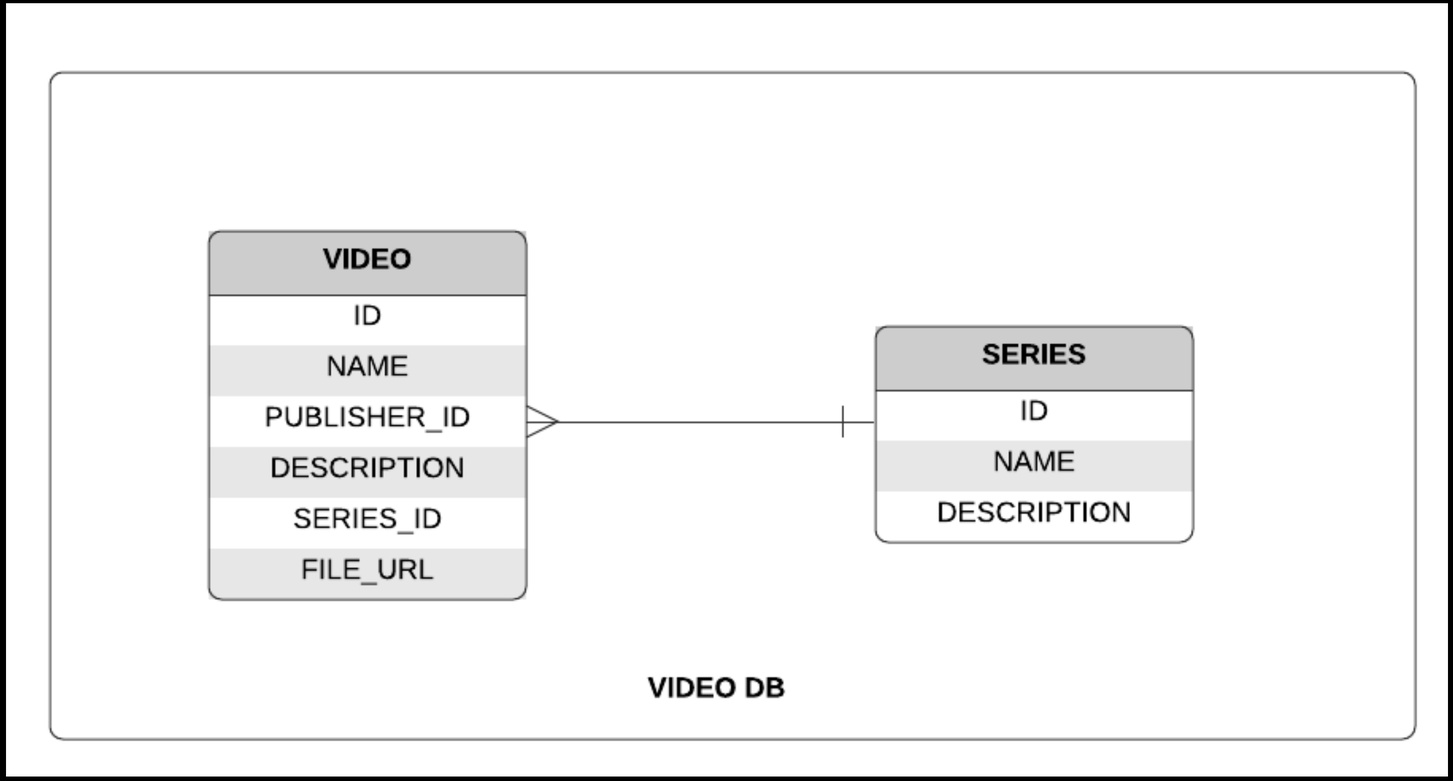
hotstar

Data Models



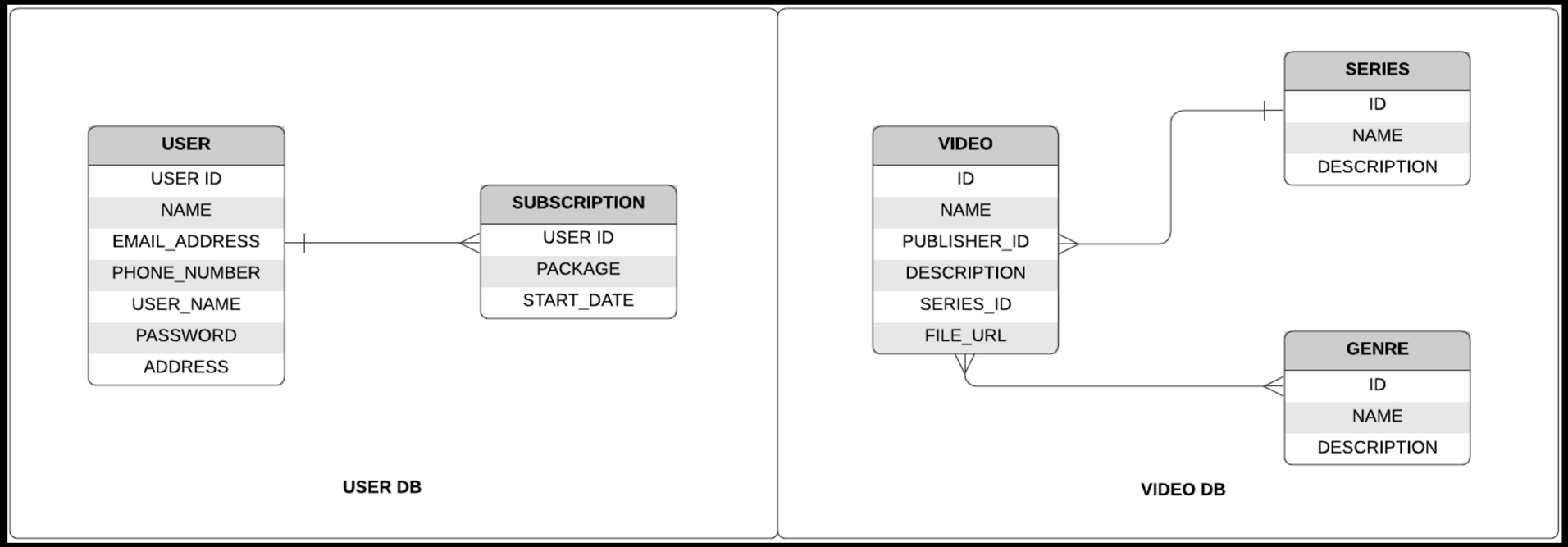
hotstar

Data Models

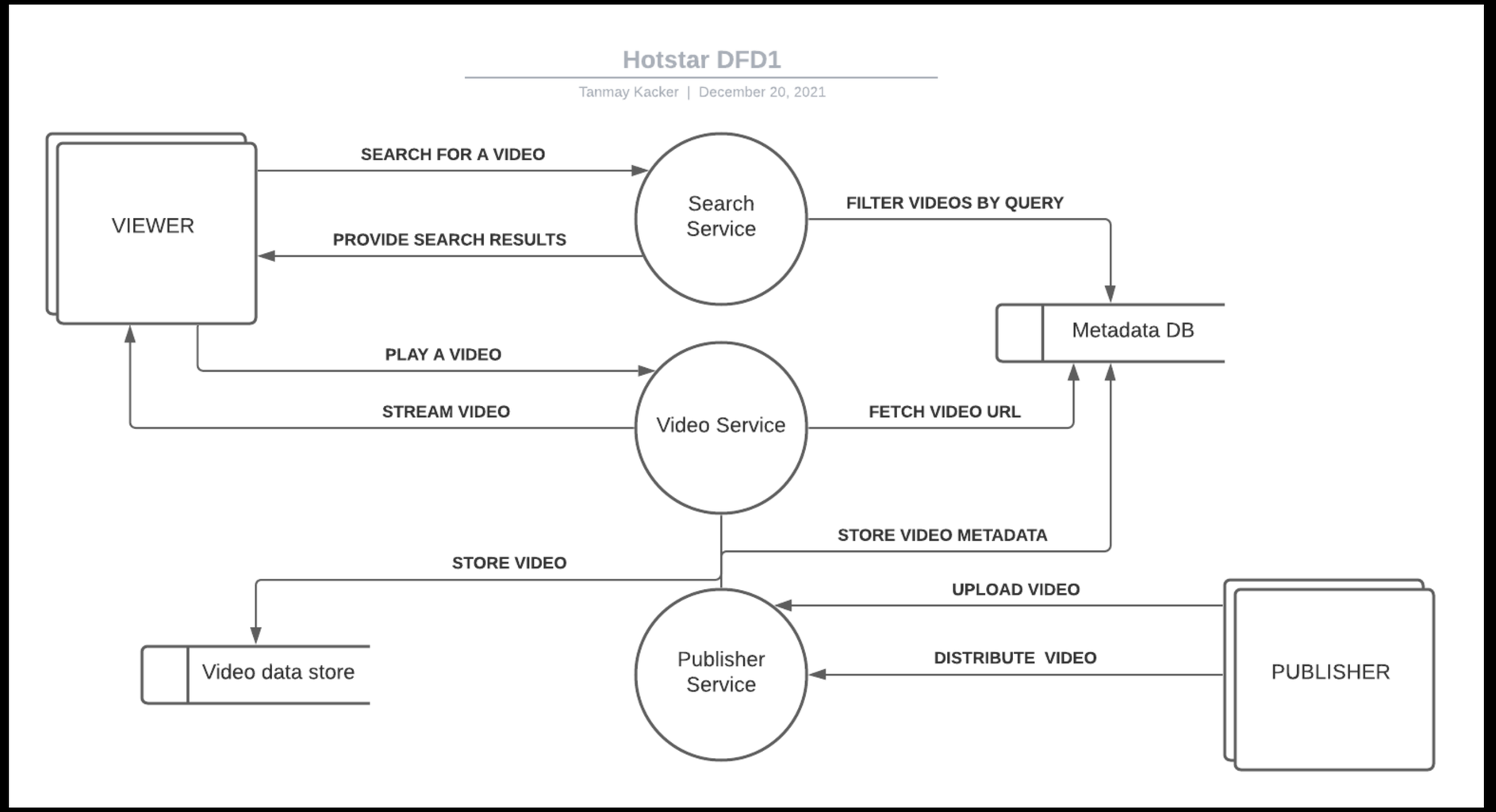


hotstar

Data Models



hotstar


Hotstar DFD1
Tanmay Kacker | December 20, 2021
VIEWER
SEARCH FOR A VIDEO
PROVIDE SEARCH RESULTS
Search Service
FILTER VIDEOS BY QUERY
Metadata DB
PLAY A VIDEO
STREAM VIDEO
Video Service
FETCH VIDEO URL
STORE VIDEO METADATA
STORE VIDEO
UPLOAD VIDEO
Video data store
Publisher Service
DISTRIBUTE VIDEO
PUBLISHER

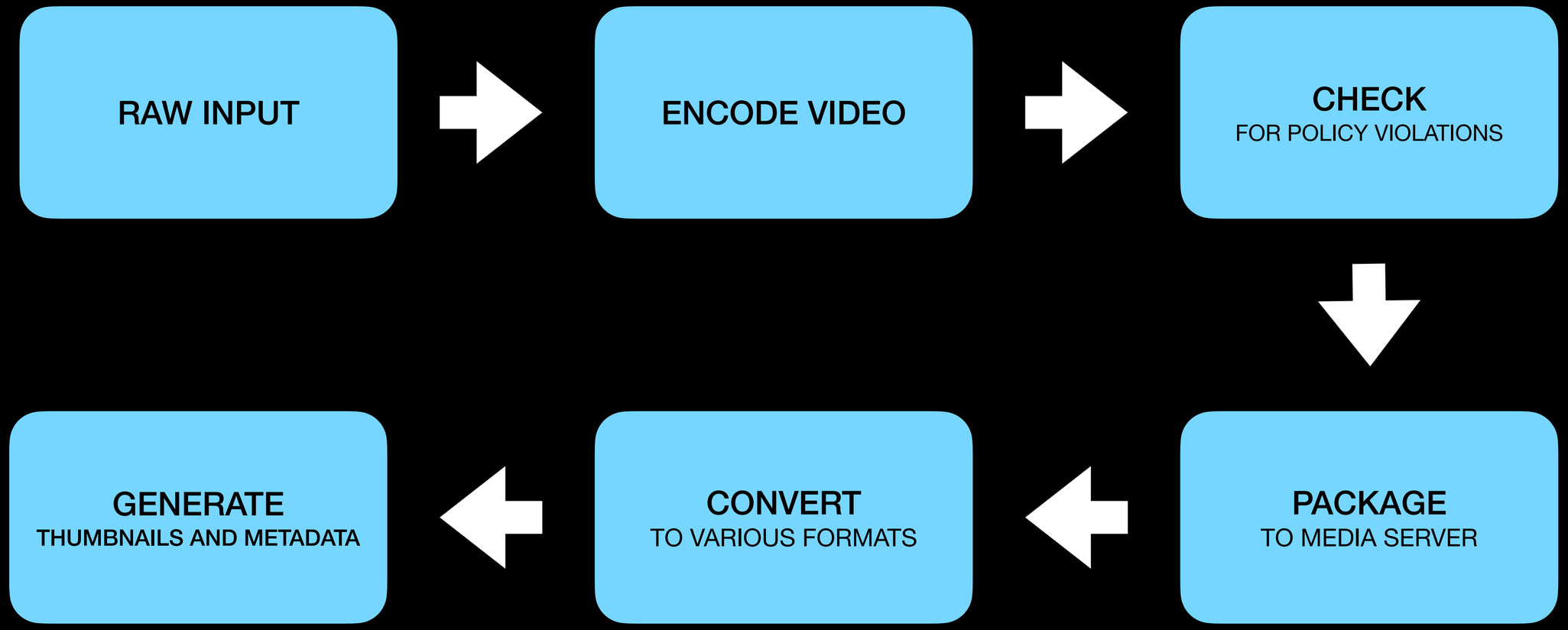
hotstar

Sample Microservices

- Account service (username, password)
- Search service (video name, genre, language, series name)
- Video service (video ID)
- Publisher service (name, series, genre, actors, market, labels, restrictions)

hotstar

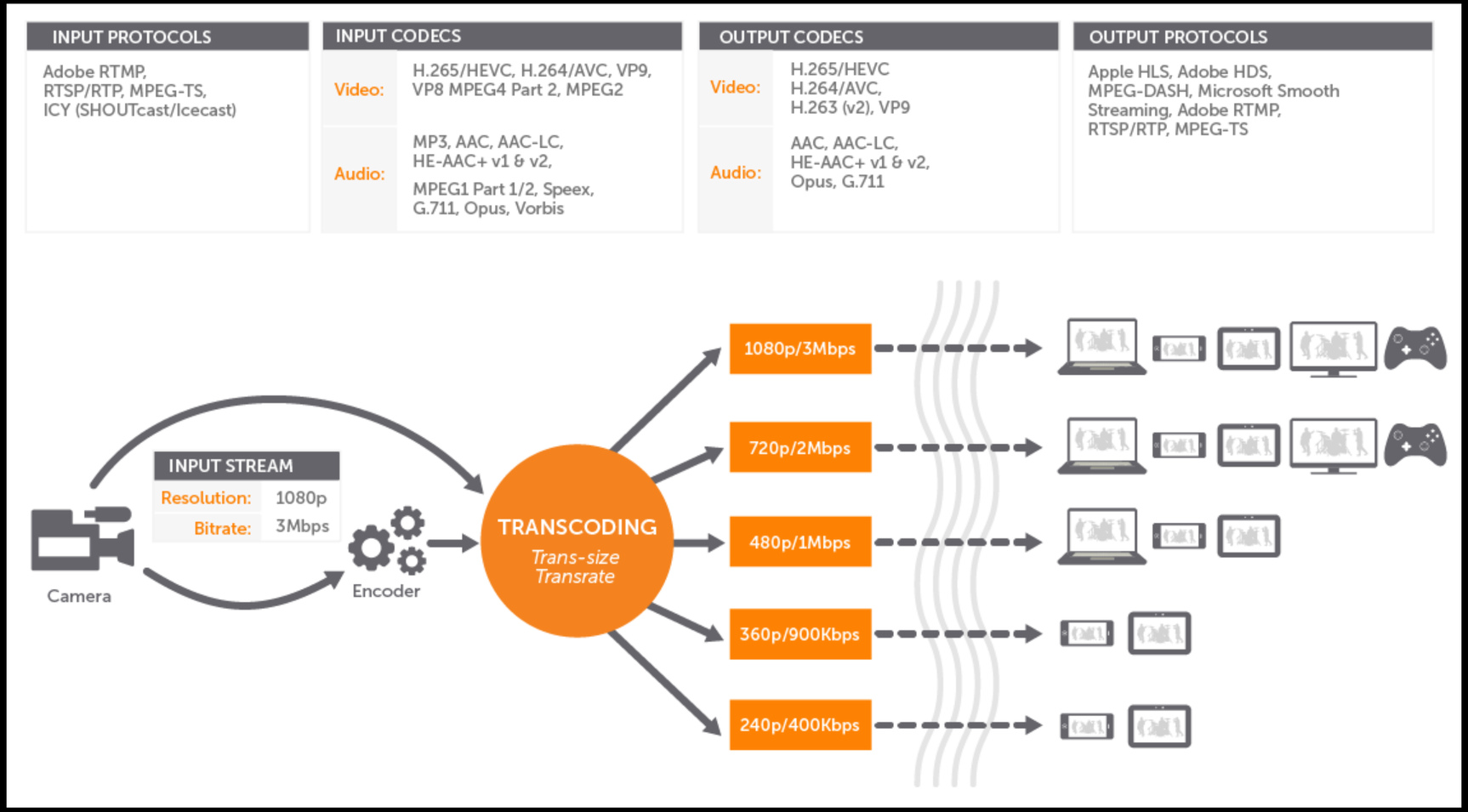
Media pipeline



hotstar

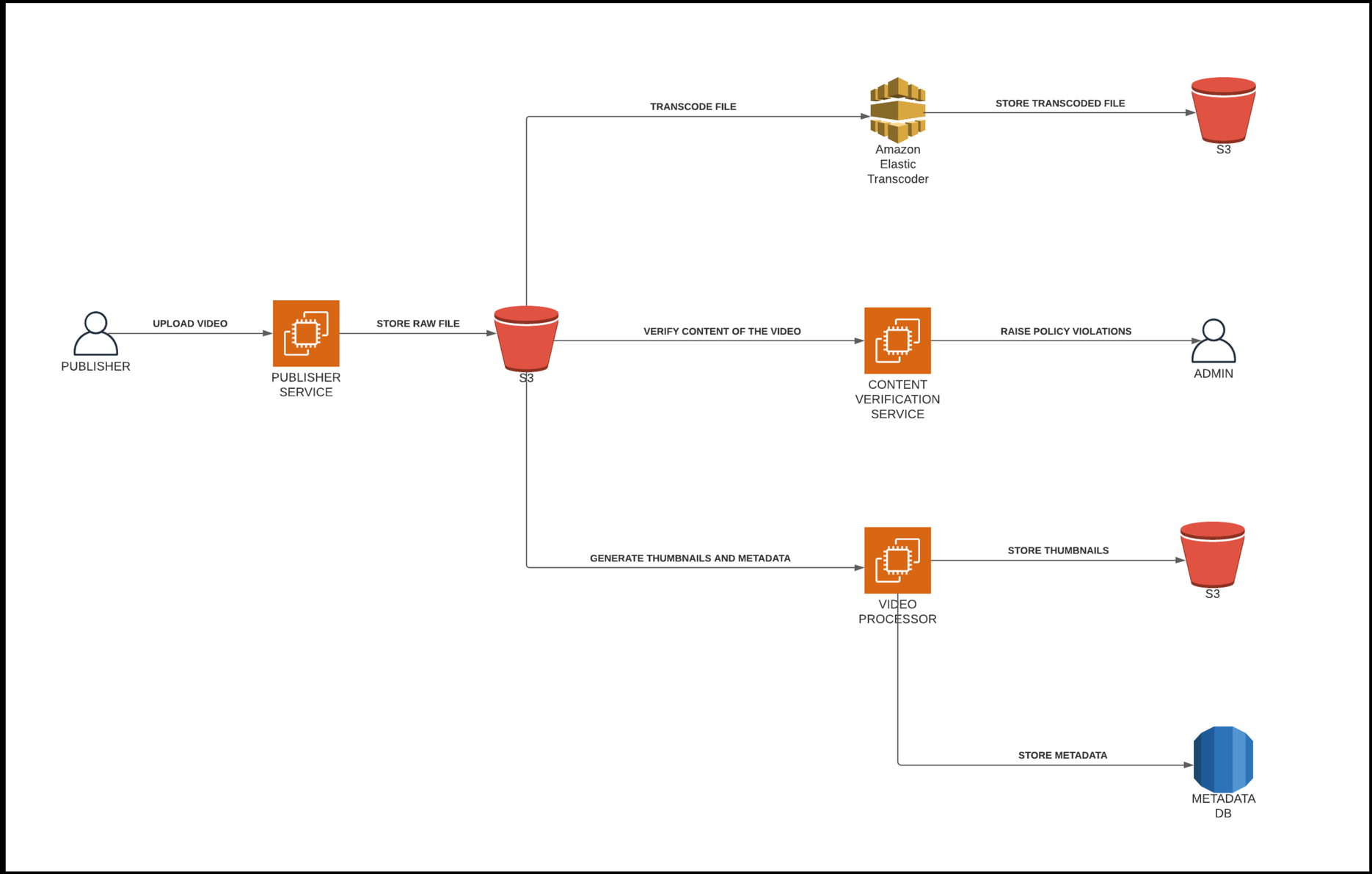
Transcoding

INPUT PROTOCOLS	INPUT CODECS		OUTPUT CODECS		OUTPUT PROTOCOLS
Adobe RTMP, RTSP/RTP, MPEG-TS, ICY (SHOUTcast/Icecast)	Video:	H.265/HEVC, H.264/AVC, VP9, VP8 MPEG4 Part 2, MPEG2	Video:	H.265/HEVC H.264/AVC, H.263 (v2), VP9	Apple HLS, Adobe HDS, MPEG-DASH, Microsoft Smooth Streaming, Adobe RTMP, RTSP/RTP, MPEG-TS
	Audio:	MP3, AAC, AAC-LC, HE-AAC+ v1 & v2, MPEG1 Part 1/2, Speex, G.711, Opus, Vorbis	Audio:	AAC, AAC-LC, HE-AAC+ v1 & v2, Opus, G.711	



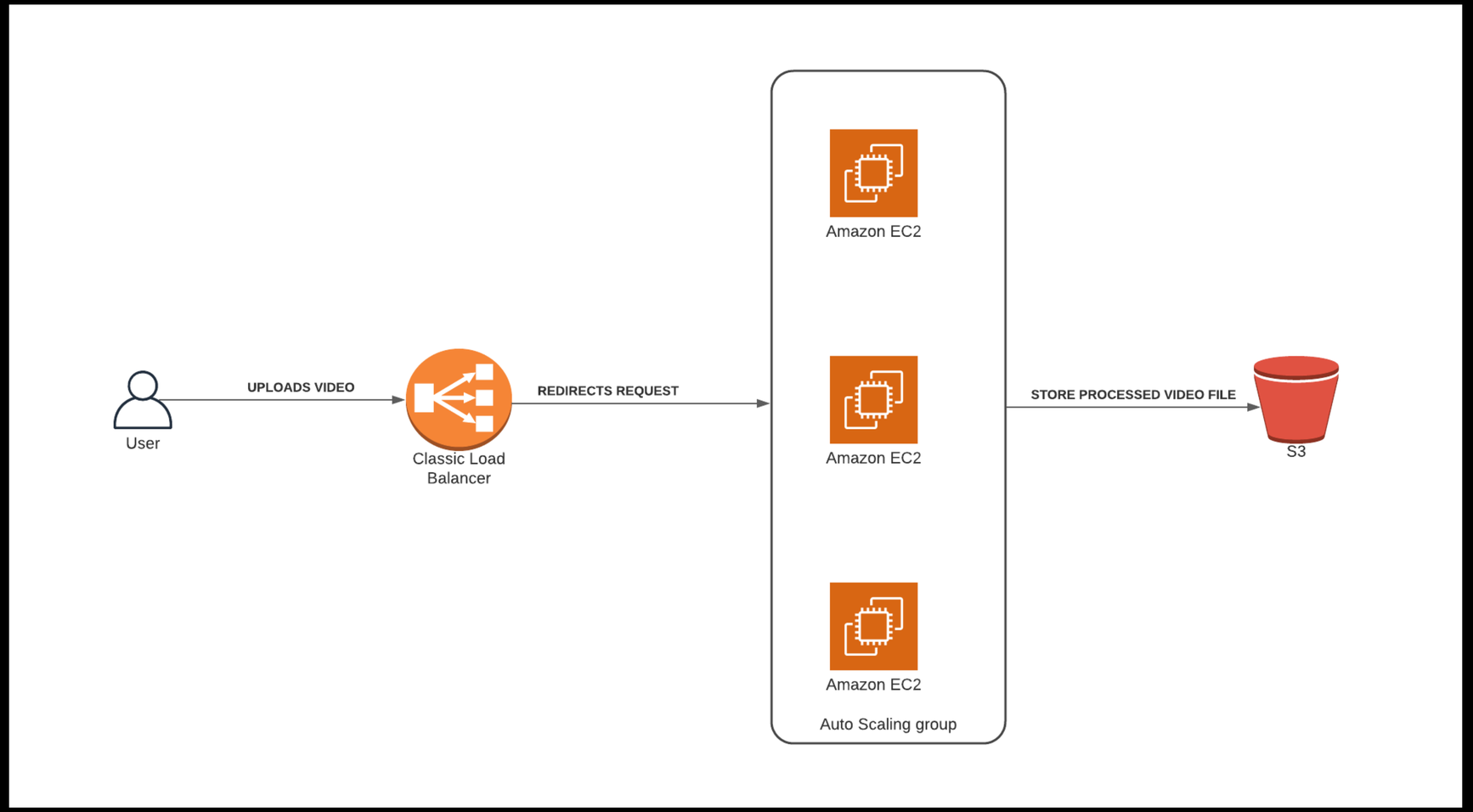
hotstar

Media pipeline



hotstar

Basic



hotstar

Basic - Problems

- Scalability - What happens if number of requests increases?
- Fault tolerance - What happens if a processing step fails?
- Latency (Blocking the sender) - The client has to wait until the last step has been processed.

hotstar

Scalability - Problems and Solutions

- Massive increase in the number of request
- What happens if a user uploads a file greater than the system’s limit?
- Horizontal vs vertical scaling
- Chunking or splitting of the file

hotstar

Fault tolerance - Problems and solutions

- Unrelated failures can cause the whole process to fail
- Intermittent and network failures
- Client has to repeat the process again for each failure
- Retryable vs non-retryable operations
- Multithreading

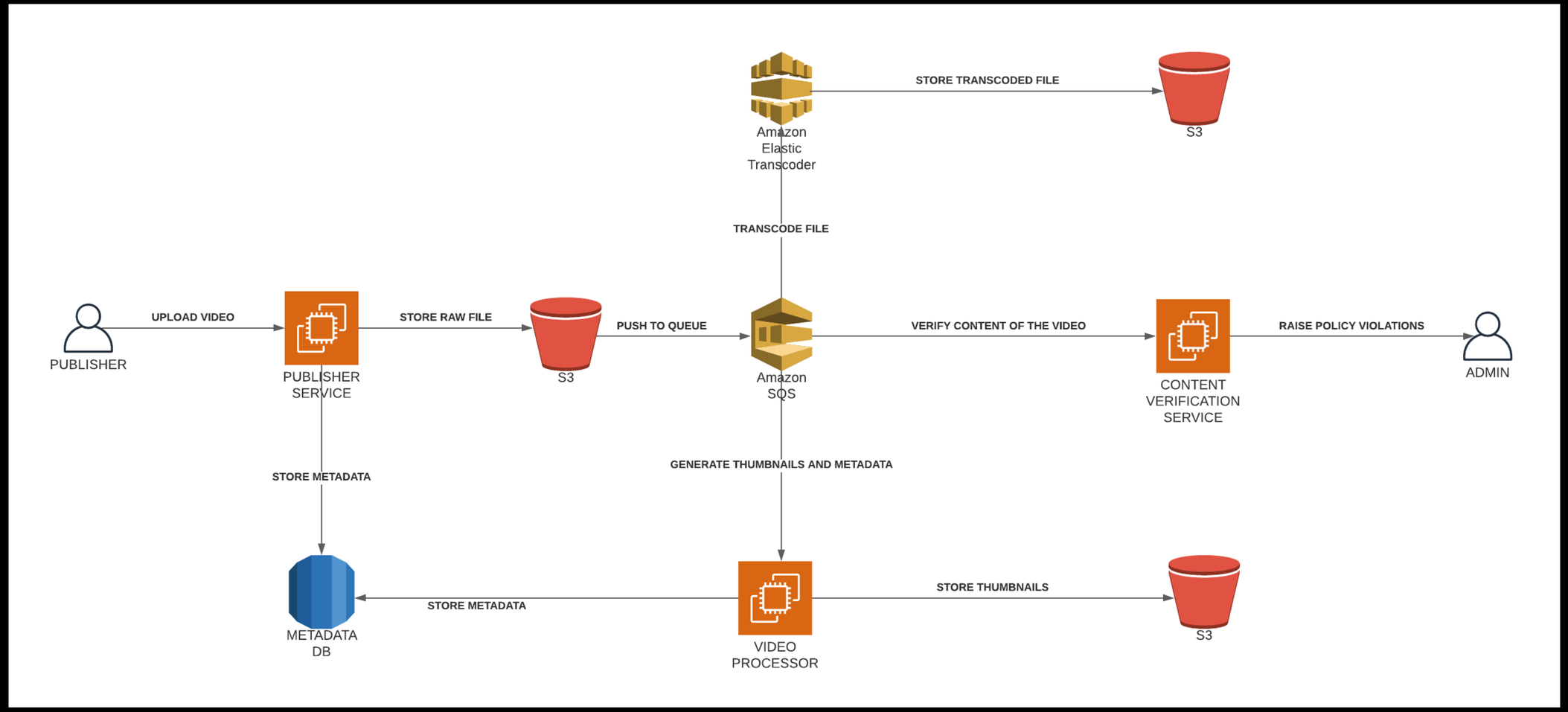
hotstar

Latency - Problems and solutions

- Client has to wait for the whole lifecycle to complete
- Bad user experience
- Client has to manually retry in cases of failure
- Reduce latency
- Parallelise operations

hotstar

Asynchronous system



hotstar

Improvements

- Scalability - Messages are picked up according to the system’s capacity
- Fault tolerance - Retry mechanisms
- Latency (Blocking the sender) - Eagerly return success message to sender

hotstar

Asynchronous system - Problems

- Complex - Additions of a queue and asynchronous paths add complexity to overall design.
- Reporting - Client is informed of success as soon as the request is processed. Failure has to be reported after closing the loop.
- Bandwidth intensive - The whole file has to transferred to the server which adds significant latency to the ingestion process
- Resource intensive - Due to dedicated workers, instances can be over-provisioned.

hotstar

Bandwith - Solutions

- Chunking or splitting of the video into parts
- Upload of video to S3 by client
- Request contains S3 link instead of the whole file

hotstar

Asynchronous system - Solutions

- Workflow management systems -
  - Airflow
  - Camel
  - Flink
  - Jenkins

Further reading

- Data flow diagrams
- Streaming protocols
- Thumbnail generation at Netflix
- Writing a DAG workflow management tool from scratch
- There is nothing as a CA system
- You can’t sacrifice partition tolerance
- Synchronous vs asynchronous systems

References

- Hotstar: Re-Architecting Apps For Scale
- Netflix: System Design
- Hotstar: Cloud based video streaming
- Scaling Hotstar to 50M

hotstar

Server to Viewer

- Size of a video file - Video files are much larger in size.
- Varying conditions - Network conditions can vary between users and even for a single user across the span of a video.
- High QPS - 20M concurrent connections.

hotstar

Solutions - Streaming

- Streaming over HTTP
- HLS vs DASH
- Adaptive Streaming
- Bitrates listed in a manifest file
- Segments listed in a sub-manifest file

hotstar

HLS - Manifest File

```
#EXTM3U
#EXT-X-INDEPENDENT-SEGMENTS
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=708416,SUBTITLES="subtitles",RESOLUTION=480x270,FRAME-RATE=29.97,CODECS="avc1.77.30,mp4a.40.5"
03/playlist.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=193587,SUBTITLES="subtitles",RESOLUTION=320x180,FRAME-RATE=7.49,CODECS="avc1.77.30,mp4a.40.5"
01/playlist.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=383667,SUBTITLES="subtitles",RESOLUTION=480x270,FRAME-RATE=14.98,CODECS="avc1.77.30,mp4a.40.5"
02/playlist.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=1236924,SUBTITLES="subtitles",RESOLUTION=592x336,FRAME-RATE=29.97,CODECS="avc1.77.30,mp4a.40.5"
04/playlist.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=1767886,SUBTITLES="subtitles",RESOLUTION=768x432,FRAME-RATE=29.97,CODECS="avc1.77.30,mp4a.40.5"
05/playlist.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=2617635,SUBTITLES="subtitles",RESOLUTION=880x496,FRAME-RATE=29.97,CODECS="avc1.77.30,mp4a.40.5"
06/playlist.m3u8
#EXT-X-STREAM-INF:PROGRAM-ID=1,BANDWIDTH=3639229,SUBTITLES="subtitles",RESOLUTION=1024x576,FRAME-RATE=29.97,CODECS="avc1.4d001F,mp4a.40.5"
07/playlist.m3u8
#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subtitles",NAME="cc1",DEFAULT=YES,FORCED=NO,URI="cc1/playlist.m3u8",LANGUAGE="en"
#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subtitles",NAME="cc3",URI="cc3/playlist.m3u8",LANGUAGE="en"
#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subtitles",NAME="Service1",URI="Service1/playlist.m3u8",LANGUAGE="en"
#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subtitles",NAME="Service2",URI="Service2/playlist.m3u8",LANGUAGE="en"
#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subtitles",NAME="Service3",URI="Service3/playlist.m3u8",LANGUAGE="en"
#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subtitles",NAME="Service4",URI="Service4/playlist.m3u8",LANGUAGE="en"
#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subtitles",NAME="Service5",URI="Service5/playlist.m3u8",LANGUAGE="en"
#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subtitles",NAME="Service6",URI="Service6/playlist.m3u8",LANGUAGE="en"
```

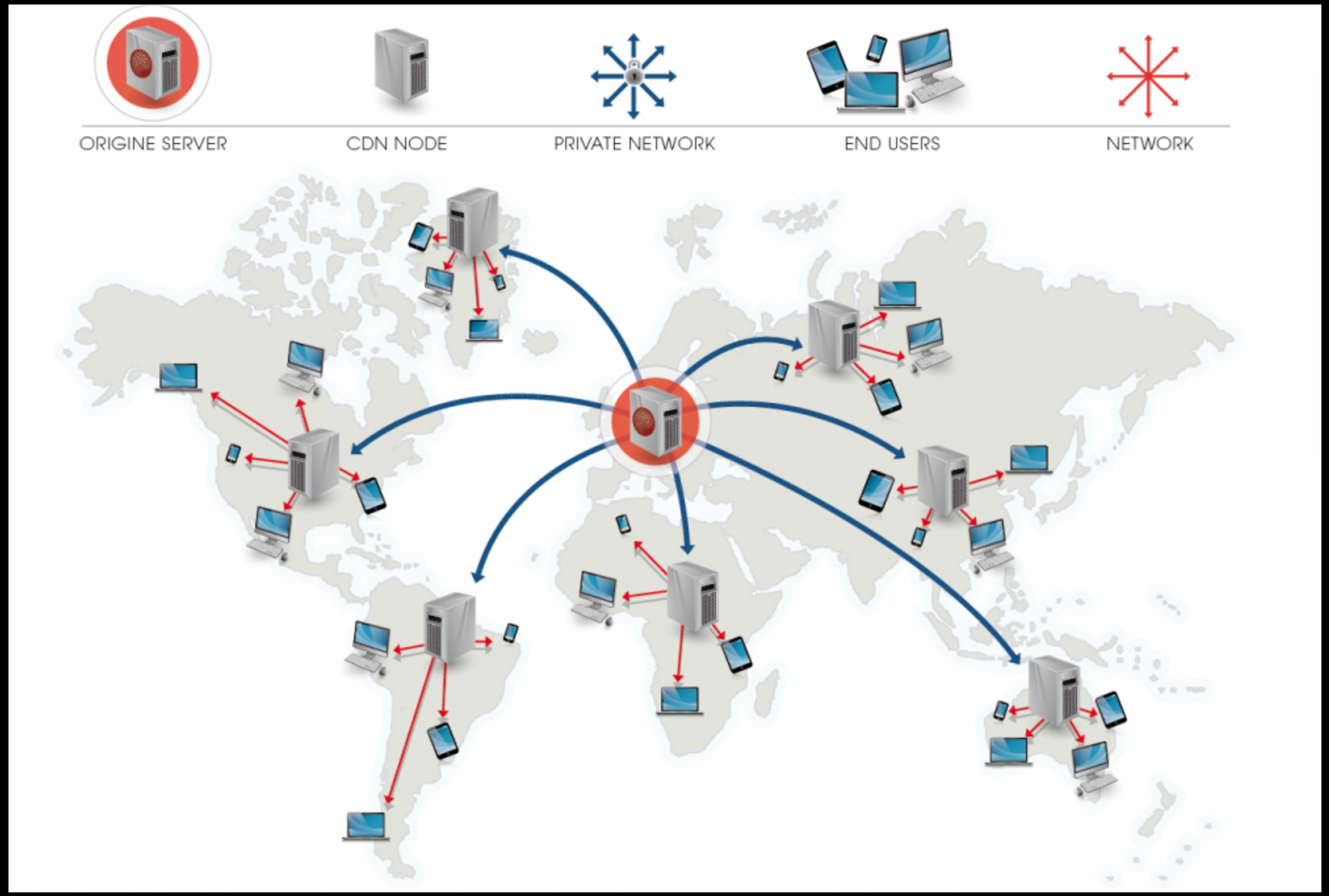
hotstar

High QPS - Problems and Solutions

- Large number of requests for one resource
- All requests are served through application servers
- Replication
- CDNs
- Data population strategies

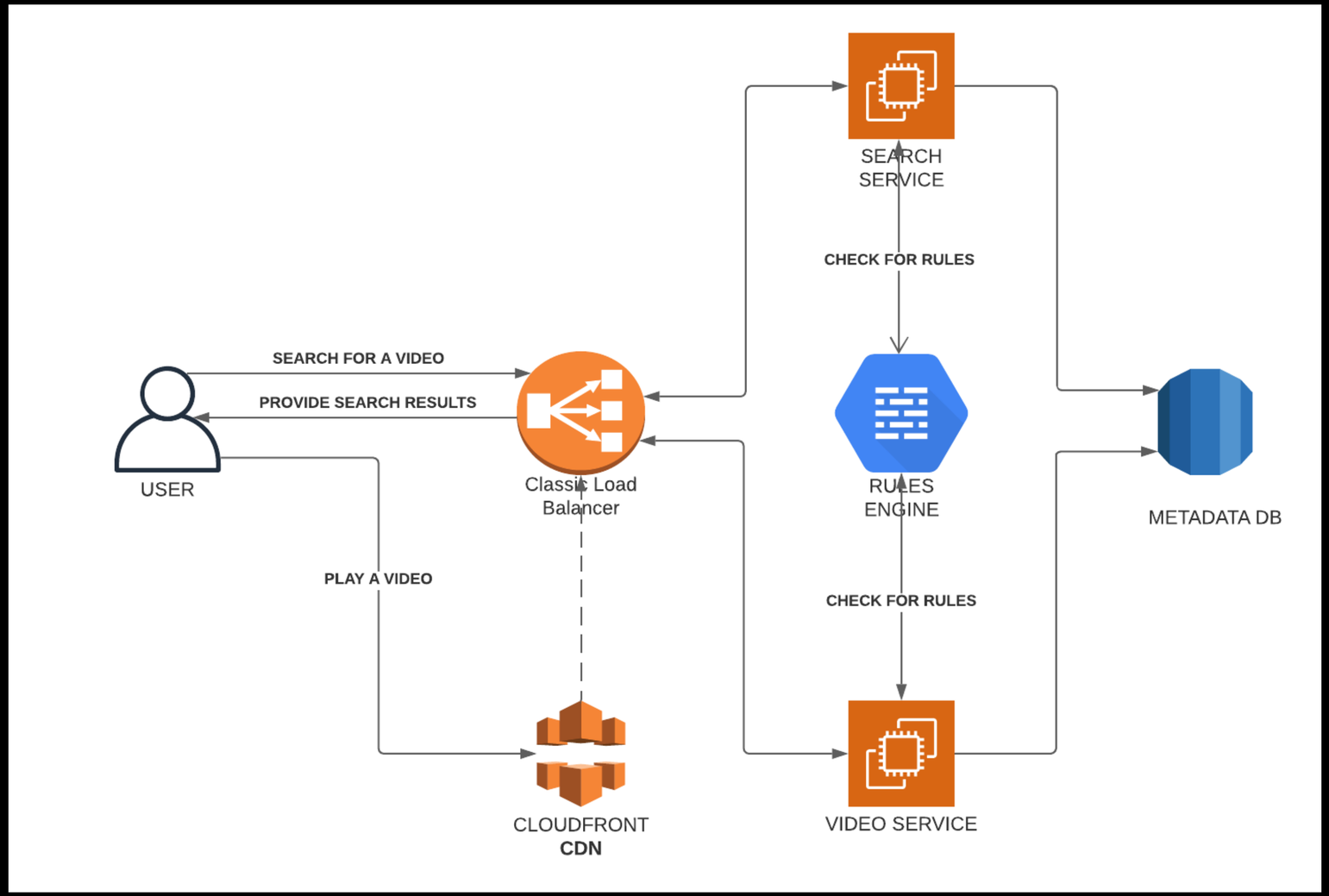
hotstar

CDNs



hotstar

Server to Viewer



Further reading

- How video works
- Video streaming
- What are CDNs?
- Understanding Netflix and CDNs
- CDN security: Through random URLs
- CDN security: Through access tokens

References

- Understanding HLS
- Netflix and CDNs

Thank You!

tanmay.kacker@scaler.com