

Video ingestion pipeline

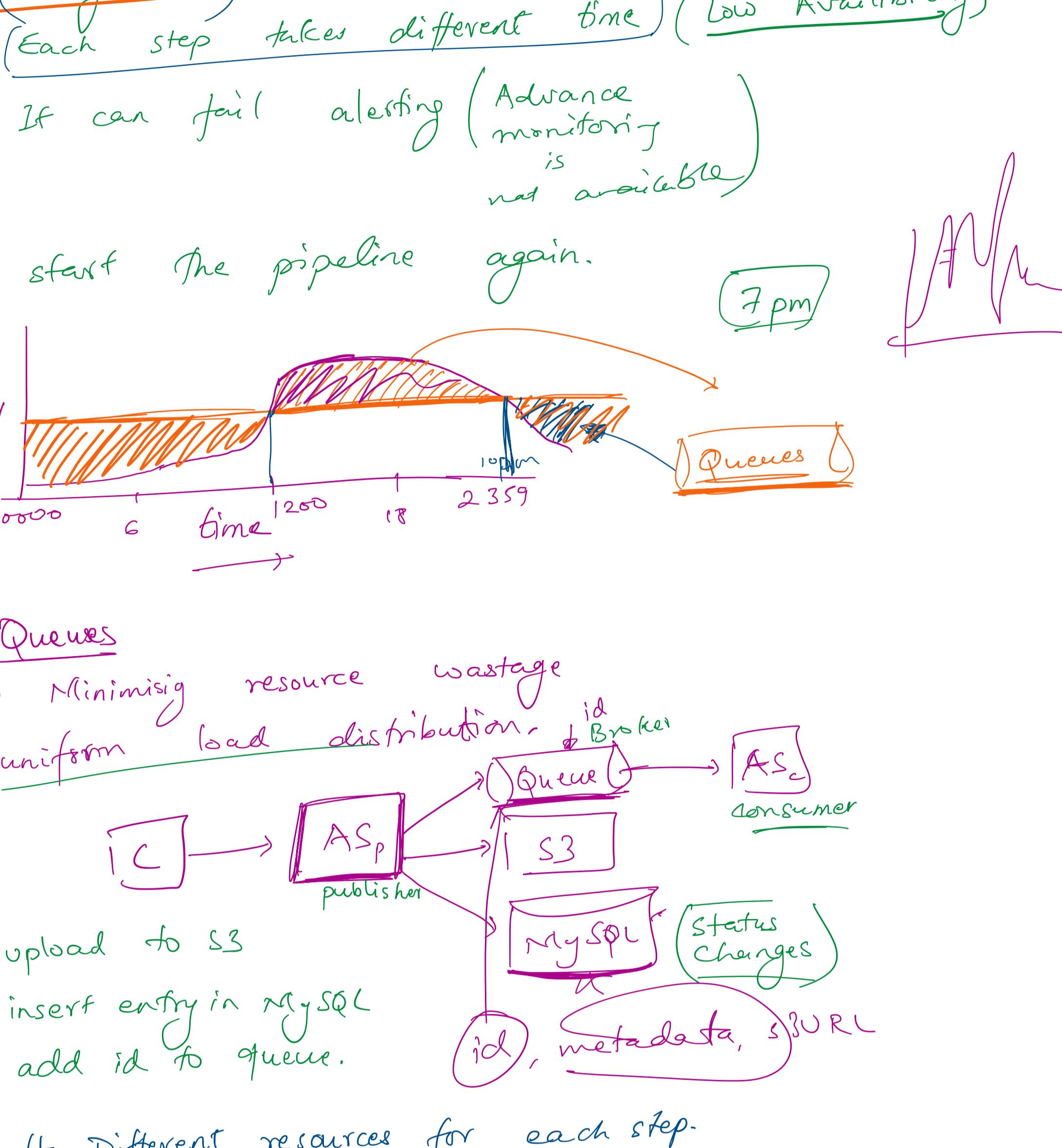
## ① Requirements

- \* any signed user should be able to upload videos.
- \* filter / search a content
- \* generating preview / thumbnails
- \* copyright check
- \* region specific laws
- \* convert to mp4 format
- \* 1080p, 720p, 480, 320, 240p
- \* segmentation
- \* -----

Segmentation → 5min

② Scale Estimation →  
③ Design schema / APIs →

## ① Initial Design

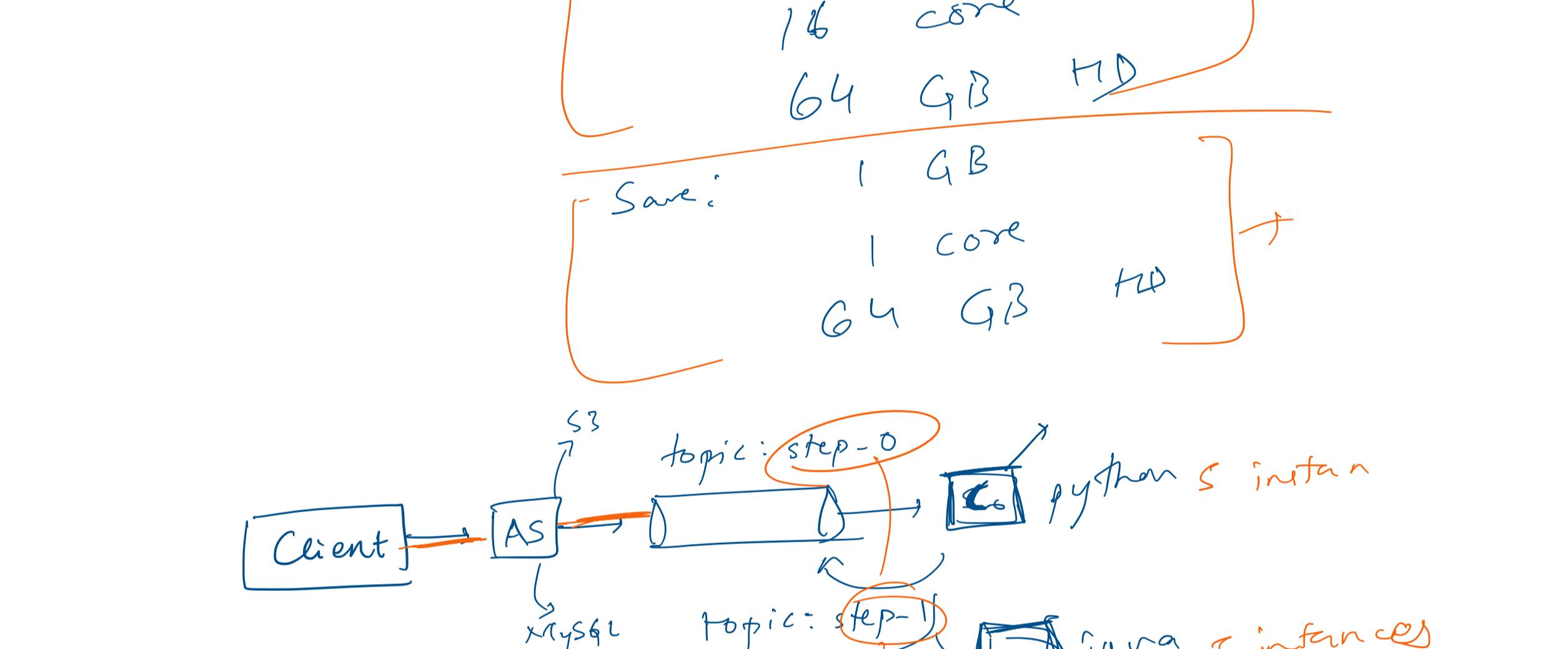


- ① Bulky code  
② Each step takes different time (Low Availability)  
③ If can fail alerting (Advance monitoring is not available)

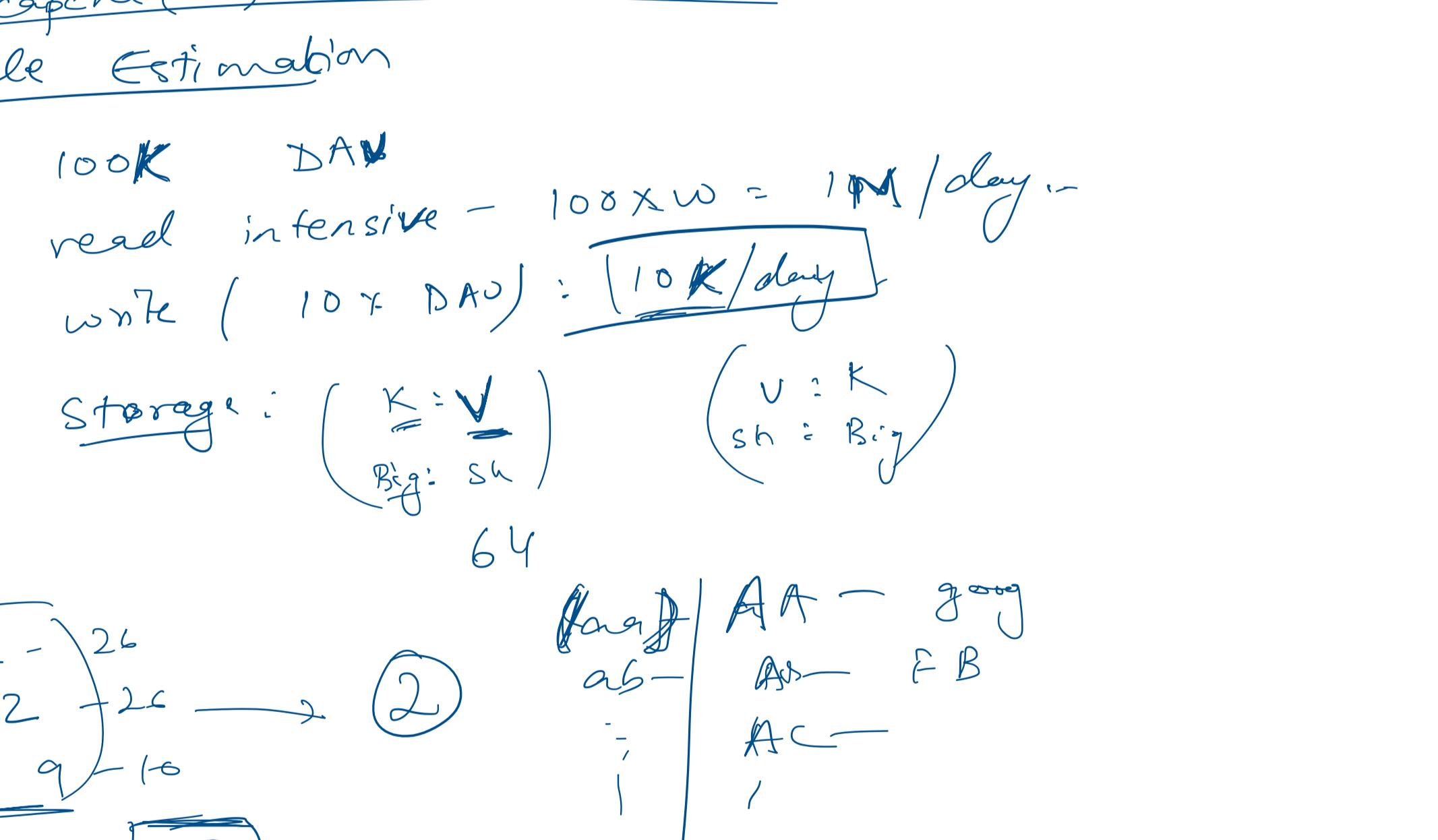
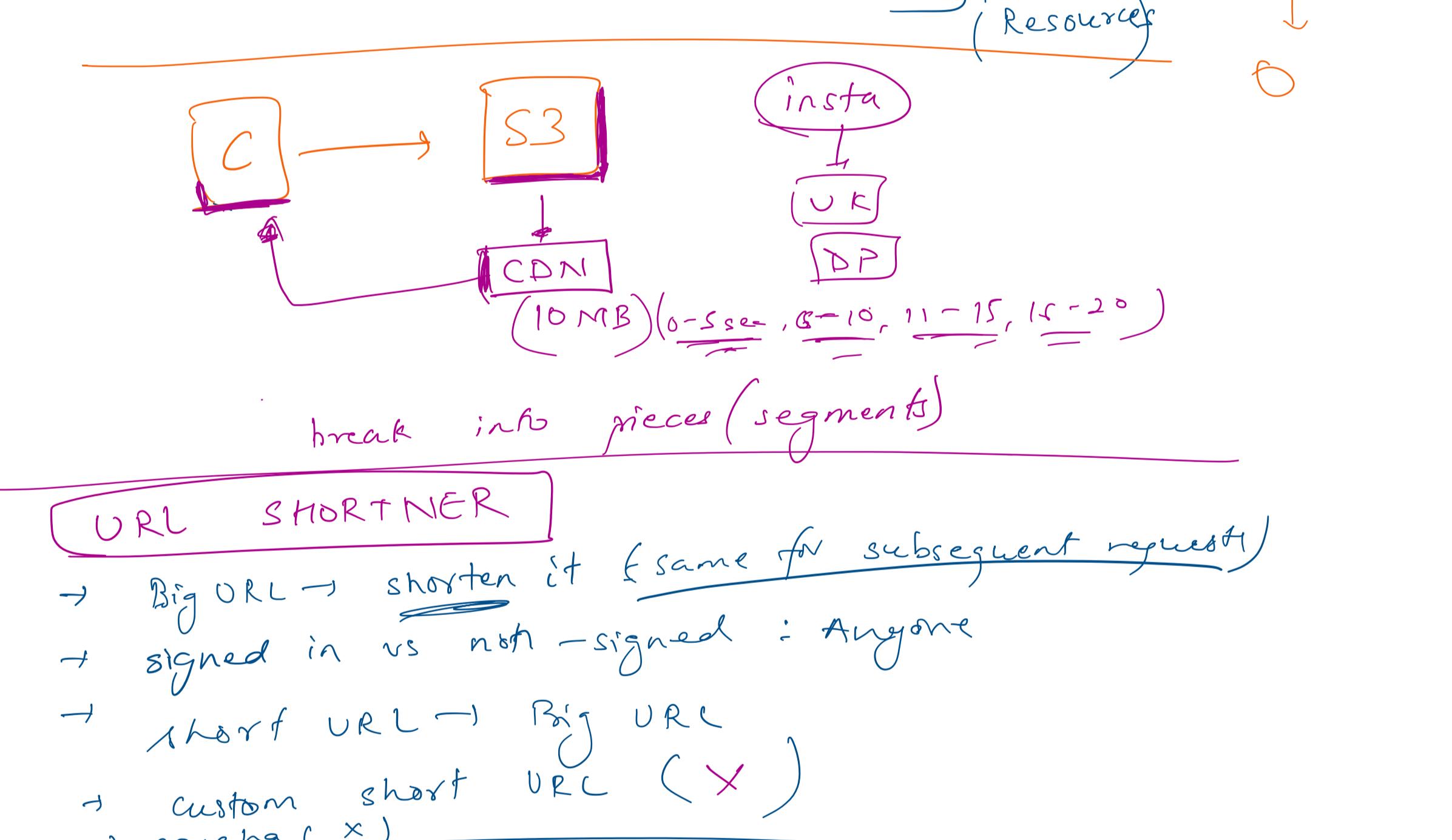
- ④ start the pipeline again.

7pm

JAN



# Different resources for each step.

URL SHORTNER

- Big URL → shorten it (same for subsequent requests)
- signed in vs not-signed : Anyone
- short URL → Big URL
- custom short URL (X)
- captcha (X)

## ② Scale Estimation

- 100K DAU
- read intensive -  $100 \times 100 = 10K$  / day
- write (10% DAU) :  $\underline{10K/day}$

$$\text{Storage: } \begin{cases} K = V \\ \text{Big: } sh \\ \text{Small: } sh \end{cases} \quad \begin{cases} V = K \\ sh = Big \end{cases}$$

## ① char :

AA - goog  
AB - FB  
AC -

1 2 22 -

32

md5()

32 chars

(a, b, c, d, e)

⇒ delete in a heap.

(a, b, c, d, e)

(AAA...A) → [1M] ⇒ AAAA

sharding?



get / short-url

post / long-url ⇒ (short-url)