Making Systems Faster: Distributed Video Processing

COS 418/518: (Advanced) Distributed Systems
Lecture 19

Mike Freedman & Wyatt Lloyd

[Grey slides from Qi Huang's SOSP 2017 Talk]

Distributed Video Processing Outline

- Motivation for video processing
  - (How streaming video works)
- Legacy design
- SVE design
- Why SVE is faster than legacy

SVE: Distributed Video Processing at Facebook Scale

Qi Huang

Petchean Ang, Peter Knowles, Tomasz Nykiel, Iaroslav Tverdokhlib, Amit Yajurvedi, Paul Dapolito IV, Xifan Yan, Maxim Bykov, Chuen Liang, Mohit Talwar, Abhishek Mathur, Sachin Kulkarni, Matthew Burke, Wyatt Lloyd

Facebook, University of Southern California, Cornell, Princeton

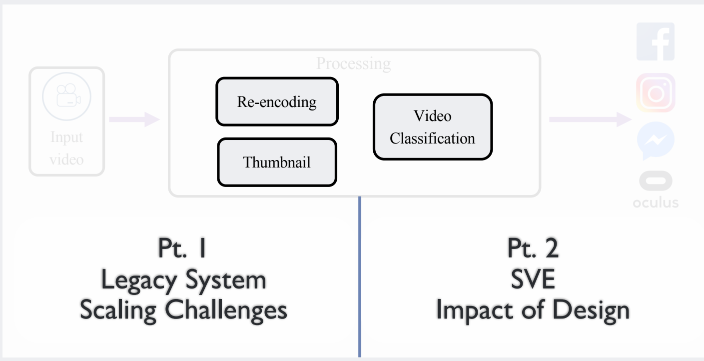
Video is growing across Facebook

oculus

- FB: **500M** users watch **100M hours** video daily (Mar. 16)
- Instagram: **250M** daily active users for stories (Jun. 17)
- All: **many tens of millions** of daily uploads, **3X** NYE spike

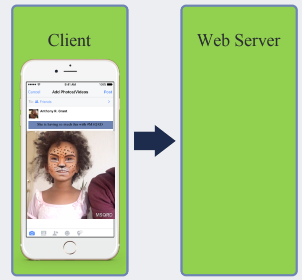
01

Processing is diverse and demanding



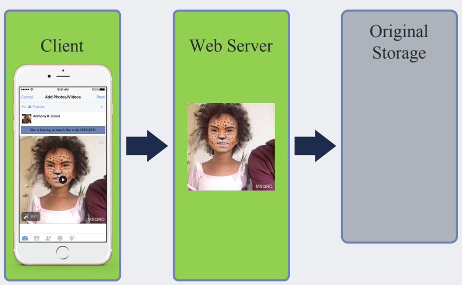
02

Legacy: upload video file to web server



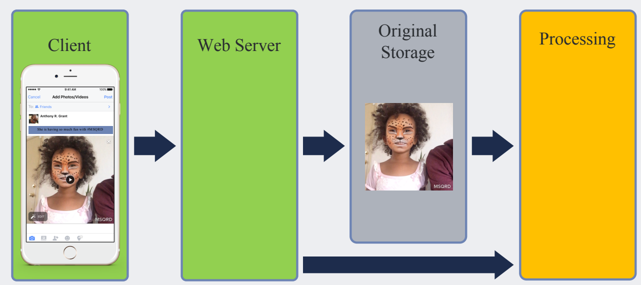
03

Legacy: preserve original for reliability

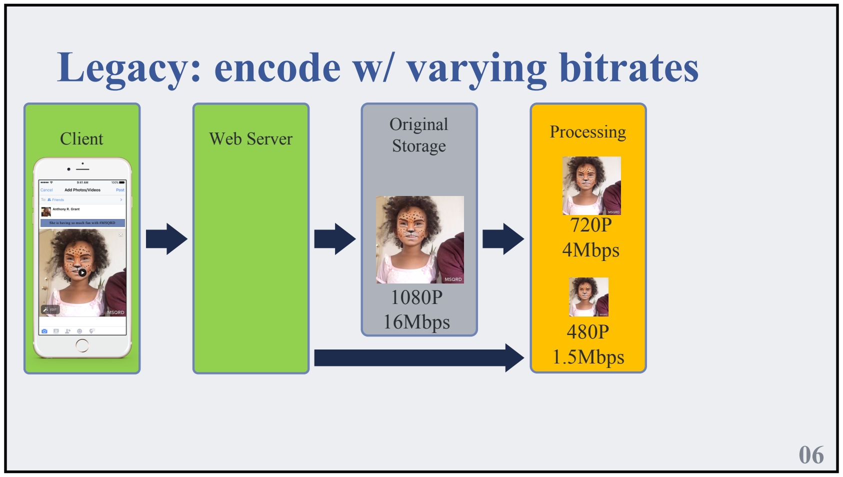


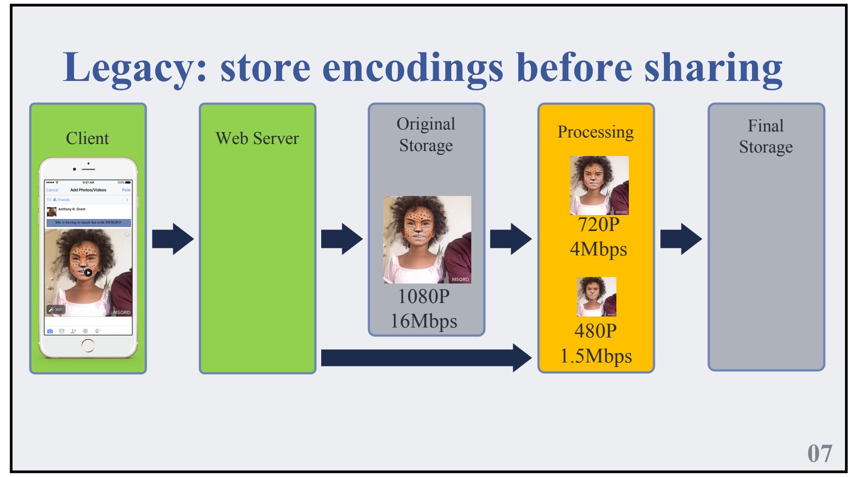
04

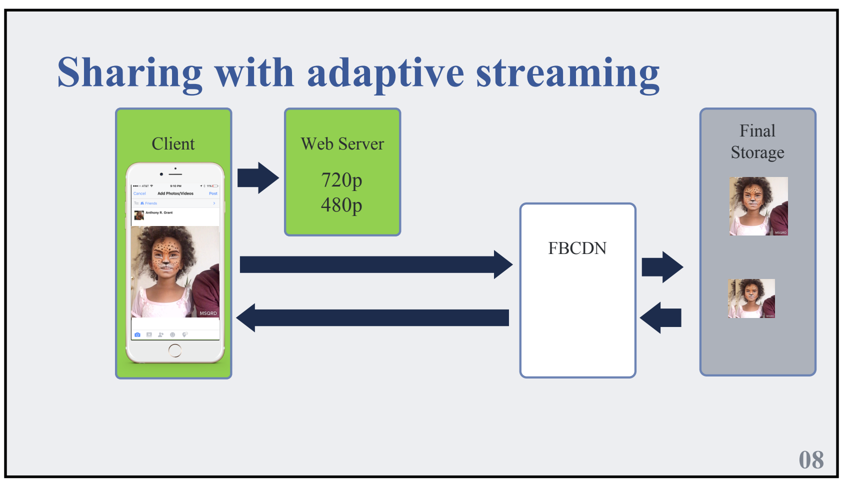
Legacy: process after upload completes

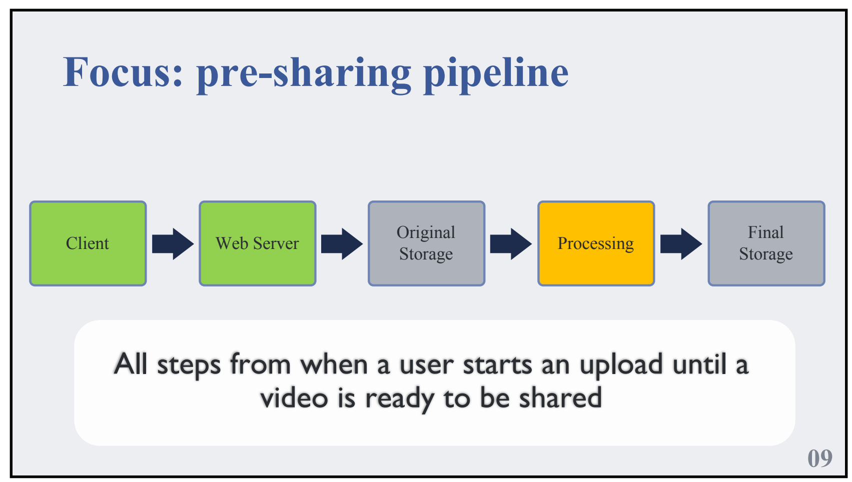


05


Legacy: encode w/ varying bitrates
Client
Web Server
Original Storage
1080P 16Mbps
Processing
720P 4Mbps
480P 1.5Mbps
06

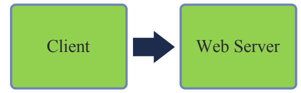

Legacy: store encodings before sharing
Client
Web Server
Original Storage
1080P 16Mbps
Processing
720P 4Mbps
480P 1.5Mbps
Final Storage
07


Sharing with adaptive streaming
Client
Web Server
720p
480p
FBCDN
Final Storage
08


Focus: pre-sharing pipeline
Client
Web Server
Original Storage
Processing
Final Storage
All steps from when a user starts an upload until a video is ready to be shared
09

How Long Does This Take? (Latency)

How Long Does This Take? (Latency)



<u>1 MB Video</u> ≈ 1 secs
8 Mbps link

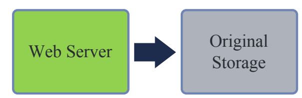
<u>16 MB Video</u> ≈ 16 secs
1 Mbps link

SVE paper stats:

<u>Video Size</u>	
≤1MB	10% of uploads over 10 seconds
3-10MB	50% of uploads over 10 seconds
300MB -1GB	50% of uploads over 9 minutes

How Long Does This Take? (Latency)

How Long Does This Take? (Latency)

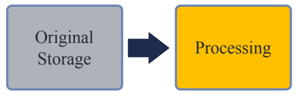


(pipelined with uploading)

SVE paper stats:

median	200 ms
90%	650 ms
99%	900 ms

How Long Does This Take? (Latency)



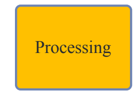
SVE paper stats:

10% of all video take $\geq$ 1.3 s

Proportional to video size:

Most videos over 100 MB take over 6 seconds

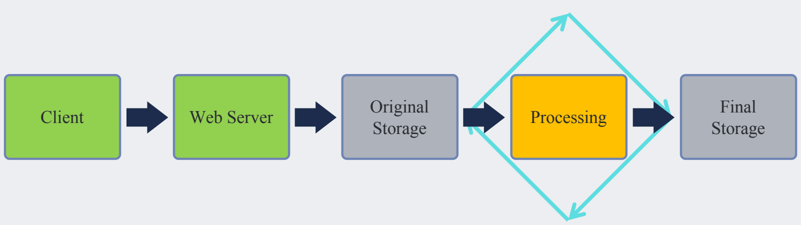
How Long Does This Take? (Latency)



SVE paper stats:

Video Size	
1-3MB	20% take over 10 seconds
100-300MB	50% take over 1 minute
>1GB	23% take over 10 minutes

Serial pipeline leads to slow processing



10

Let's Make This Faster!

Talk to classmates about how!

Speedy: harness parallelism

Users can share videos quickly

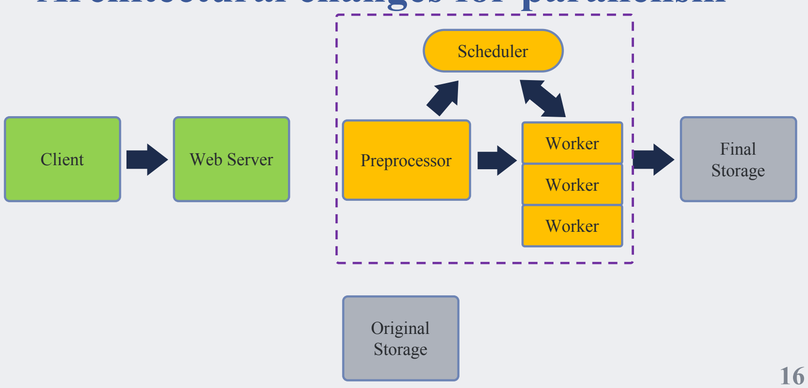
- Overlap fault tolerance and processing
- Overlap upload and processing
- Parallel processing

14

Architectural changes for parallelism

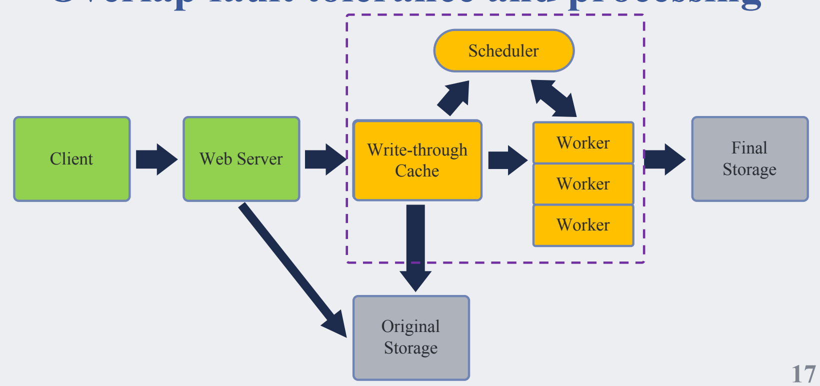
15

Architectural changes for parallelism

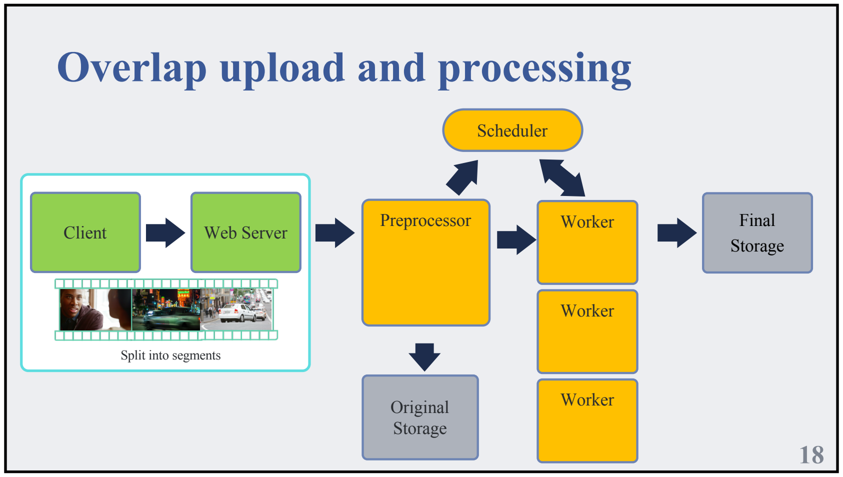


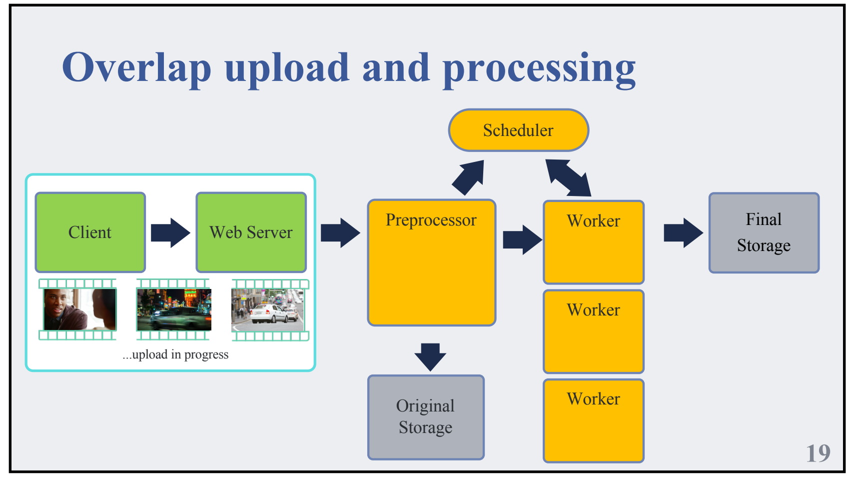
16

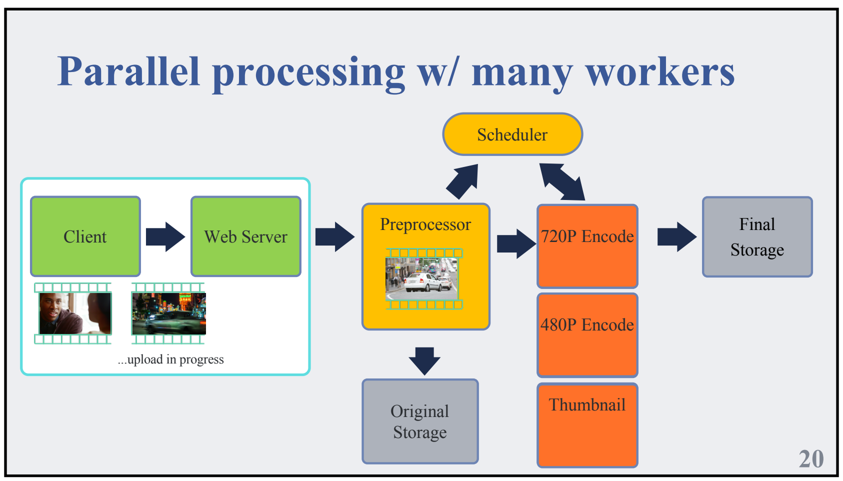
Overlap fault tolerance and processing

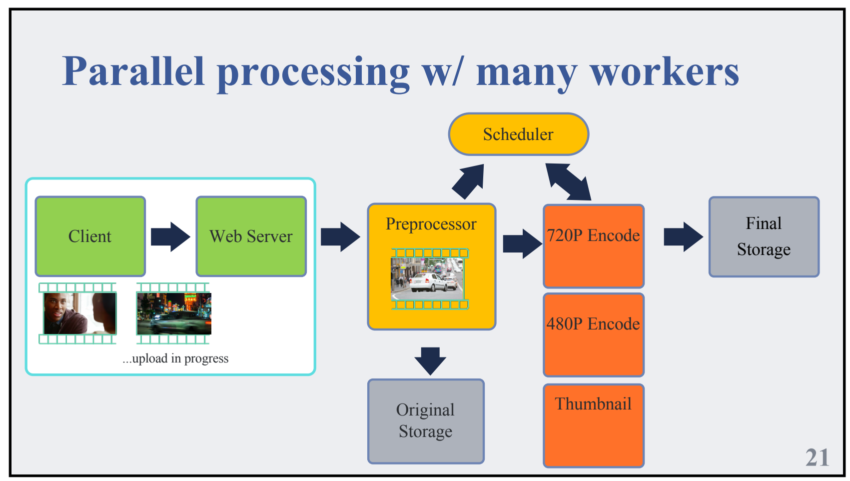


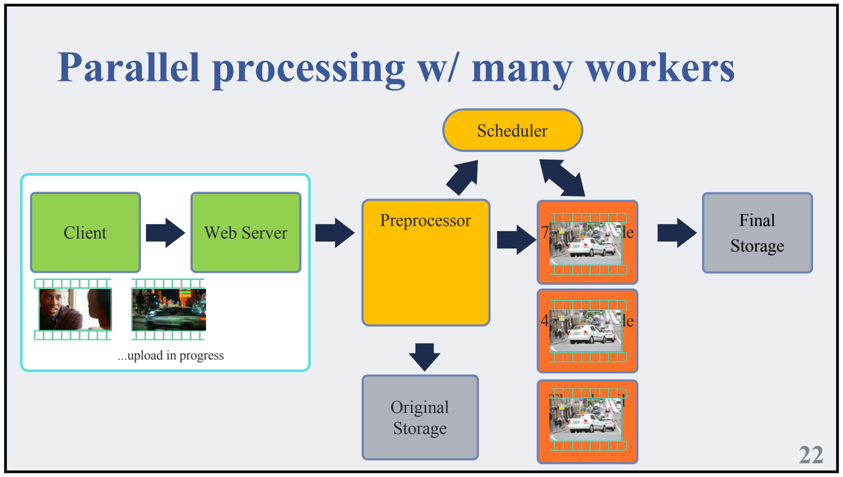
17

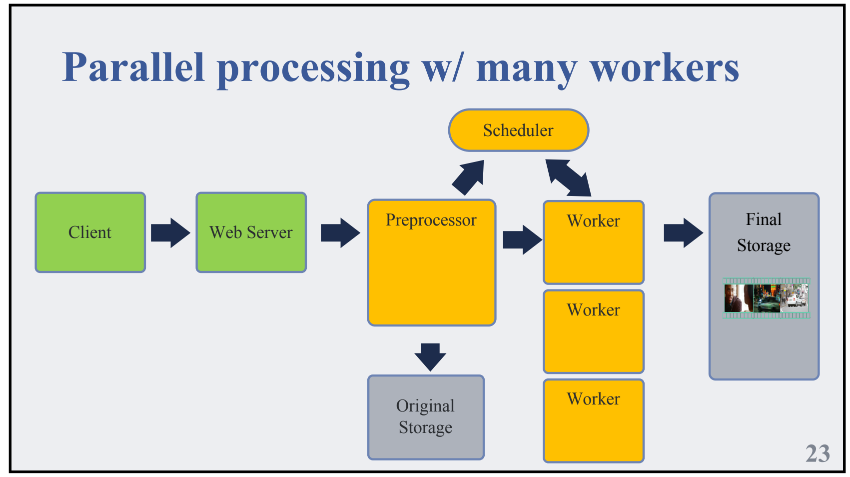

Overlap upload and processing
Scheduler
Client
Web Server
Preprocessor
Worker
Worker
Worker
Final Storage
Original Storage
Split into segments
18

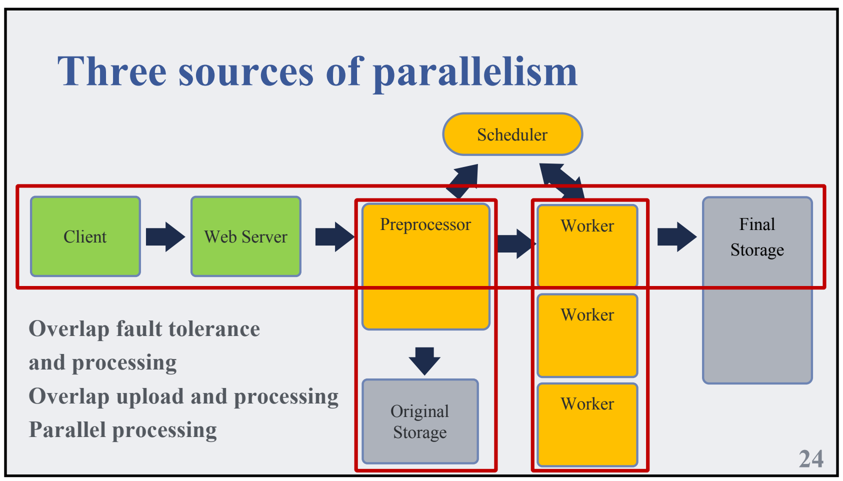

Overlap upload and processing
Scheduler
Client
Web Server
Preprocessor
Worker
Worker
Worker
Final Storage
Original Storage
...upload in progress
19

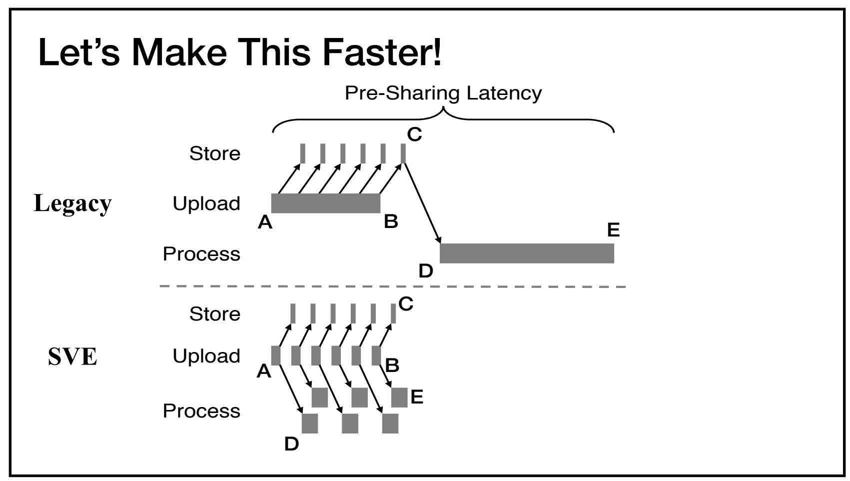

Parallel processing w/ many workers
Scheduler
Client
Web Server
Preprocessor
720P Encode
480P Encode
Thumbnail
Final Storage
Original Storage
...upload in progress
20


Parallel processing w/ many workers
Scheduler
Client
Web Server
Preprocessor
720P Encode
480P Encode
Thumbnail
Final Storage
Original Storage
...upload in progress
21

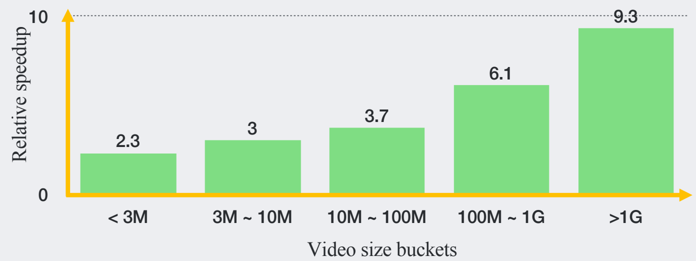

Parallel processing w/ many workers
Scheduler
Client
Web Server
Preprocessor
Final Storage
Original Storage
...upload in progress
22


Parallel processing w/ many workers
Scheduler
Client
Web Server
Preprocessor
Worker
Worker
Worker
Final Storage
Original Storage
23


Three sources of parallelism
Scheduler
Client
Web Server
Preprocessor
Worker
Worker
Worker
Final Storage
Original Storage
Overlap fault tolerance and processing
Overlap upload and processing
Parallel processing
24

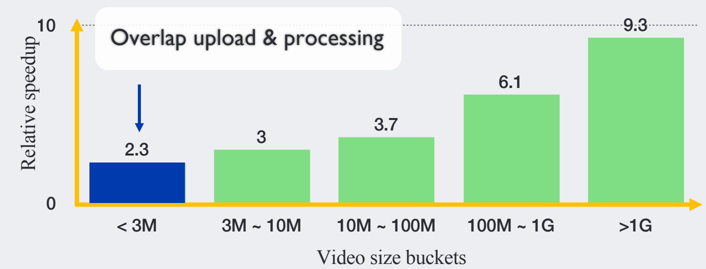

Let's Make This Faster!
Pre-Sharing Latency
Legacy
Store
Upload
Process
SVE
Store
Upload
Process
A
B
C
D
E

Results: 2.3x ~ 9.3x speedup



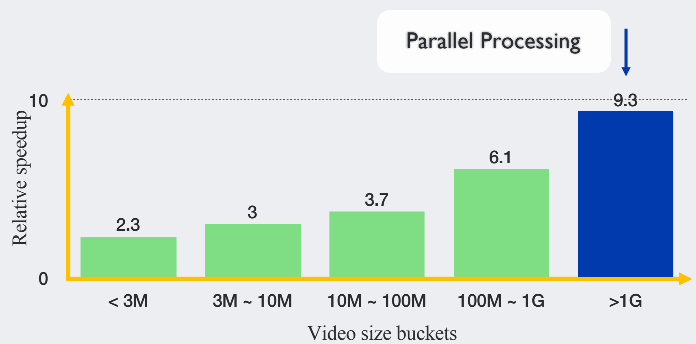
25

Results: 2.3x ~ 9.3x speedup



26

Results: 2.3x ~ 9.3x speedup



27

Summary

- Motivation for video processing
  - (How streaming video works)
- Legacy design – Serial processing was slow
- SVE design – Three sources of parallelism make SVE faster
  - Overlap upload and processing
  - Overlap fault tolerance and processing
  - Parallel processing